Computer animation of deformable bodies

Samoylenko, Igor

Award date:
2002

Awarding institution:
University of Bath

Computer Animation of Deformable Bodies

submitted by Igor Samoylenko

for the degree of PhD

of the University of Bath

2002

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The problem of modelling elastic media was rigorously addressed more then a hundred years ago with the development of the classical theory of elasticity. However, the theory in its linear form has serious practical limitations when applied directly to modelling highly elastic solids. We have shown that it is the local rotational misalignment of the reference frame and the deforming body at a point that introduces errors in the evaluation of the stress tensor and therefore of the elastic force at the point. This misalignment problem seriously limits the range of deformations the linear classical theory can handle. Indeed, for any deformation that results in the rotational misalignments of the reference frame and the deforming solid at some points of the solid, the elastic forces at those points will be evaluated with error.

In the theory of FlexyMatter we have added another step into the classical path of the evaluation of the elastic force: before calculating the displacement field in the neighbourhood of a point, the deforming solid and its reference frame should be aligned. After that the displacement field can be calculated, followed by the evaluation of all other related quantities (strain, stress and finally the elastic force). By additionally performing this alignment we have shown that this leads to a simplification of the expression for the strain tensor. Additionally, the first covariant derivative of the displacement field is equal to the ordinary first partial derivative. This simplifies further the classical expression for the strain tensor, as well as the expressions for the stress tensor and the elastic force.

Our theory of FlexyMatter is a direct extension of the linear classical theory of elasticity and as such is mathematically rigorous. It does not compromise on the mathematical rigour to produce an easy to use modelling technique. It is based on the linear classical theory of elasticity and therefore benefits from its long established status and the wealth of the accumulated experience.

We have shown that our theory of FlexyMatter satisfies all the criteria for a successful modelling theory. It is rigorous, complete, flexible and versatile.

Acknowledgements

I would like to thank my supervisors Dr Claire Willis and Dr Derek Paddon for their enthusiasm, continuous support and excellent supervision during this research. They have always been there for me with their help and advice. Doing my research with them has been a very interesting experience for me, something that I will have good memories about for the rest of my life. If I had to do it all over again, I would not want to do it with anybody else. Without them this research would have been impossible.

These acknowledgements would, of course, be totally incomplete without mentioning my wife Joy. She has always been my source of inspiration, my guiding light that always shined for me and gave me the strength to carry on and bring this research to its conclusion. I would like to thank her from the bottom of my heart for always being there for me, for believing in me, for supporting me during my often-unsociable working hours, for understanding its importance for me and in general for being the wonderful person that she is! If without Claire and Derek this research would have been academically impossible, without Joy, it would have been simply impossible.

I would also like to express my gratitude to my family in Ukraine: my mother, our family friend Viktor Nikolaevich Lisovoi, my brother Viktor and my sister Anya, as well as my new family in Thailand: Joy's mother Oranush, Joy's father Suvit and her sister Shirley. I would like to thank them all for their support and always believing in me.

Table of Contents

Table of Figures

1. Introduction

The classical theory of elasticity has served the scientific and engineering community for more than a century. Some problems can be solved within the framework of classical elasticity but others may have to be modified or even simplified to enable a solution to be obtained. Possibly the most difficult class of problems that defies effective solution are those that exhibit large deformations for even small amounts of strain. So the deformations of membranes, rods and other thin objects have been studied by using very specialised techniques. Solid but floppy materials fall outside of these techniques and no effective method for their solution has been reported in the literature.

In this thesis our aim is to develop a theory for modelling a wide range of elastic deformations: from almost rigid to highly elastic. The intended application of such a theory is computer animation. Therefore, it should be complete (in the sense that it should provide all means necessary to implement a practical modelling system), easy to use, versatile and provide for fast and stable practical implementations.

Physically based modelling of flexible bodies in the area of computer graphics has become a very popular subject for applied research. A number of theoretical approaches have been proposed [7, 12, 18, 27, 38]. A wide range of practical areas of research and problems associated with them received attention, including:

- Areas of research: numerical approximations of the equations, their stability, precision, and speed; collision detection and response; the problem of control of dynamic simulations; physically based surface appearance modelling.
- Various applications: modelling cloth (static, draping, on synthetic actors), modelling human-like objects, and knitting simulations.

Its popularity is not surprising. The natural desire to achieve better realism in simulations was coupled with a phenomenal growth in available computational power. In some areas of research (such as cloth modelling) the interest has also been fuelled by the industry: a system, capable ofaccurately and interactively simulating various types of cloth, their easy manipulation (incuding seaming, cutting etc) and accurate rendering, clearly has tremendous commercial opportunities associated with it.

The existing difficulties in modelling 3D flexible solids can be best demonstrated by the most recent state-of-the-art commercial film projects: "Walking with Dinosaurs" and "Walking with Beasts" by the BBC in the UK and "Toy Story II" cartoon film from the USA.

All three films produced very impressive pictures of simulated reality; they included vivid lighting effects, realistic textures etc. But it is a known fact that very little physically based simulation was actually used and as the result they both failed to produce the kind of physically realistic and convincing reality we experience in everyday life. Cartoon characters in "Toy Story II" and dinosaurs in "Walking with Dinosaurs" were impressive but did not really have the physical qualities they would have had if their movements and interactions had been simulated according to the classical laws of mechanics, and they had the classical mechanical properties associated with any object in real life, such as mass, elasticity and rigidity, incorporated into their motion.

Our primary goal in this research is to devise a theory for modelling the behaviour of complex three-dimensional solids as they deform under the action of applied forces.

Even though the theory we develop will be designed to model the real world, it is easier to relate to a computer-simulated world instead. In our real world, we have certain preconceptions: gravity is always directed downwards and cannot be easily "turned off"; we cannot fly without proper equipment and so on. However, in a simulated world (cartoons, computer simulated reality films, etc) all these things are achievable.

Let us analyse the problem we must address in more detail. We need to deal with deformable objects in our simulated world, for example: a rubber ball thrown against the wall, a cartoon character walking on the floor and being hit by the ball. We need to account for external factors: gravity is suddenly "switched off" and objects start floating freely in the air; later it is turned back on but acting in a different direction. We need to consider possible

constraints: a cartoon character's foot is suddenly glued to the floor and would not move whatever happens to the character.

We can see that a successful theory must be able to answer the following questions:

- If we have an object (say, a rubber ball) and throw it against the wall in the given direction with the given initial velocity, how will it bounce back? How will the material (rubber) properties affect its behaviour?
- When the ball hits a cartoon character, how will the cartoon character fall down? How will the cartoon character deform as he collides with the floor?
- What is the role of the gravitational force? How can we control it: turn it on/off, change its direction and magnitude? How can an artificial force be added?
- We want to add external forces and influences, like wind, air resistance. We also want to add some constraints, such as the ability to fix an object or part of an object permanently (i.e. it does not move whatever happens around it).
- We must ask how all of the above can be achieved?

We can already see from this simple overview the kind of objects we need to deal with, the interactions they may encounter, the external conditions and influences (external forces, like gravity, wind, air resistance, various constraints) and internal conditions (object properties, like mass, material properties and so on) we need to account for. In particular:

- **The objects**: we will deal with three-dimensional deformable objects, made of some solid material. We assume the material and its properties are known or can be measured/discovered practically.
- **The interactions**: objects may interact (collide). The way two objects will interact is completely defined by their properties (material they are made of, mass and so on), the surface conditions and properties (surface friction) and external conditions (external forces). It is therefore clear that, after all of the above is specified, the theory must provide a complete description of object interactions.
- **External conditions and influences:** there will be external forces and external influences in our system. The most common example of the external force is gravity, but it can also be a variety of applied forces, for example, wind, air resistance and so

on. By external influences we mean constraints, such as "at all times one point of an object must stay fixed at the given point in the space or on another object".

Our artificial world is controlled by an animator and it is clear that our theory must provide the means for the animator to control only the factors that are directly of his/her interest: material properties (including object mass) and the conditions on their boundaries (friction), external forces (gravity, wind, air resistance and so on) and constraints. Everything else must be determined by our theory, or, in other words, everything else must be directly derivable from all the specified factors. The theory must provide a conclusive description of how to calculate object positions, velocities, deformations, interactions and so on.

1.1 Mechanical description of the modelling problem

Here we provide a detailed analysis of the modelling problem from both mechanical and physical standpoints. At the end of the section we will develop a precise mechanical formulation of the modelling problem in the form that will be used throughout this thesis.

Objects to be modelled

We will be dealing with material objects, solids, to which we attach properties, including the mass (more precisely mass distribution throughout the body), and elastic properties whose precise nature is to be defined later.

The objects are solids; we will not deal with gases or fluids, as they have fundamentally different physical properties comparing to solids. We will use the following definition of a solid.

Definition (1.1)

An object is considered to be a solid if it can sustain shear stresses (that is, the stresses that try to turn a square object into a skewed one).

This definition conforms well to our general understanding of the physical structure of solids, fluids and gases. We know that everything consists of tiny atoms, but in solids, they have well defined structure relative to each other; in fluids and gases there is no such a structure and the molecules can float around freely. This is a key difference, which has a direct impact on the physical properties of the objects.

One can see why gas and fluid are not solid, according to this definition: their molecules move chaotically, and therefore generally will not keep their original shape under applied shear forces during the experiment.

Object interactons

Our objects can interact with each other. During the interaction objects exert force on each other through their area of contact (the contact area being a part of an object's boundary). We can see that the object interaction will be fully accounted for, if we consider an arbitrarily distributed boundary force as being given at all times. By boundary force we mean the force acting through the object boundary, measured in force per unit area. When we say the boundary force is given at all times, we mean it is either prescribed (by the animator) or it can be measured (as with for example friction) at any time.

External forces

These will include:

- Gravity like forces: the forces that act on mass, and are distributed throughout the body. These forces are called volume forces (measured in force per unit volume). Gravitational force is a natural example of such a force. Besides gravity, there may also be artificial volume forces added by the animator. In fact, as we will see later, there is another volume force that appears in a body as soon as it is deformed. This is called an internal (or elastic) force and it will be considered in detail later in section 2.7.

- Boundary forces. These forces act on the body through its boundary only. They are called boundary forces (measured in force per unit area). We have already identified one possible boundary force caused by the object interactions. We can combine all these forces by assuming we have an arbitrary boundary force given at all times, and this will be the resultant of all boundary forces at each point (interaction between objects, wind, air resistance, friction and so on).

We can now see we only need to deal with two fundamentally different kinds of forces: volume and boundary forces (more precisely: with their resultants) distributed throughout the body.

Now we can provde a complete mechanical formulation of our modelling problem in the form that will be used in this thesis:

> Given a solid, its distribution of mass, material properties, initial position and velocity, the distribution of volume and boundary forces (given at all times, not just at the start of the experiment), the task is to determine the position, velocity and the deformation of the solid throughout the experiment.

1.2 History of the Physically Based Modelling over the last 20 years

The main purpose of this overview is to create a consistent picture of the main developments in the field of physically based computer modelling from the mid 1980's up to the present time. The twenty years as the definite timescale for the overview was not chosen arbitrarily. Up to approximately the middle 1980s physically based modelling in computer graphics existed mostly as a theoretical area of research mainly because the computational resources required were unavailable at that time.

The overview starts with a description of the state of affairs in the field of physically based computer modelling before the mid 1980s. It then proceeds to describe the major developments during the last 20 years from the point of view of physically based modelling.

1.2.1 The state of affairs before the mid 1980s

Once the computer became more than just a powerful calculator, and its computational capabilities reached a significant level, computer graphics has become an integral part of applied computing. Modelling the "real" world, using computer graphics, started to attract interest from the research community. Initially, animation and geometric modelling were used as the means of producing realistic scenes of the objects:

- Animation required all the individual frames to be produced by a human operator and the computer was simply used as a device to draw, store, play back and process those frames and frame sequences.

- In geometric modelling, on the other hand, geometric primitives were used to construct scenes. The operator prescribed the specific shapes, positions, velocities, trajectories etc and the computer task was simply to generate, store and playback the individual frames and their sequences.

As computer power grew, various techniques were developed to facilitate the production of realistic scenes. In the "key frame animation" method, for example, each scene was considered as a system with a number of control parameters (positions, velocities, trajectories etc). For each key frame the operator had to specify the values of those parameters and the computer would generate all the required frames in between. This, combined with the ability to use geometric primitives with textures (human generated pictures) attached to them, allowed a skilled animator to produce impressive animated sequences.

However, there existed fundamental difficulties: the production of a realistic scene required an a-priori knowledge of what exactly was going to happen at every instant. The animator was required to have a very vivid imagination to be able to generate a scene that would look "live" and realistic. This was only satisfactory in a limited number of applications.

It became evident that the only way to address these difficulties was to develop new techniques based on our knowledge of the processes in the real world. The computer's role in this case would become increasingly more important. From being just a convenient device for manipulating images it would turn into a rendering engine that would use the relevant knowledge from mathematics, mechanics, physics and other branches of science to generate images and image sequences where the scene objects would behave in a physically consistent fashion, and thus be more realistic. In other words, there is a need to simulate physical reality by computing solutions from a theory that is consistent in all aspects to the real world of solid behaviours.

Our theoretical understanding of phenomena relevant to the applied modelling of world scenes was mostly established in the 19th century. Various theories were developed, including the theory of rigid body movement, the theory of elasticity, thermodynamics and so on. They were able to produce precise equations governing the phenomena found in the real world, including rigid body movement and elasticity (and plasticity) of bodies under the action of applied forces.

There was, however, a serious obstacle in the way of applying these theories to practical modelling problems. The equations are very rarely (only in most simple cases) solvable explicitly. As the result, application of numerical methods to solve these equations was the only option. Hence, it required significant computational power to make it practically applicable and the computational power necessary for these tasks only started to become

available in the 1970s. However, the nature of numerical modelling of mathematical equations is difficult and normally demands that the equations are simplified in some way before the numerical model can be applied. In the case of equations defining elastic deformation, it requires linearisation of the equations, which has the end result of restricting the range of deformation by a very significant amount.

By mid 1980s computer technology had been given a tremendous boost. New processors were developed that were faster often by a factor of two than their predecessors every few years, with generations of processors changed within 5-7 years. In fact, that trend not only continued over the years but also dramatically increased. The lifespan of generations of processors nowadays is shrinking at an exponential rate, being currently at a level of 1-2 years.

The need for new modelling techniques coupled with rapid increases in the available computational power created favourable conditions to attract the increasing attention of the research community. As this historical overview will show, the popularity of physically based modelling would steadily increase from the mid 1980s.

1.2.2 Major developments

First published work on physically based modelling of flexible bodies: 1986

In 1986 the first work on flexible object modelling in computer graphics was published. Weil [5] was the first to apply physical laws governing the appearance of flexible objects subjected to applied forces to modelling the shape of hanging cloth suspended at a finite number of points. He used the fact that the curve described by a heavy rope hanging in a gravitational field is a (part of) catenary. The problem addressed by Weil was static, the cloth was discretised into a finite number of nodes, and the equilibrium shape was approximated by "sweeps" – cycles during which every node was moved towards its equilibrium position. The direction and magnitude of these movements were determined from the positions of the neighbouring nodes.

In the same year, Feynman [6] used a similar approach to model the appearance of hanging cloth: energy minimisation. The surface of the cloth was discretised into a collection of individual nodes. Deformation energy was introduced, a scalar function E of positions, endowing the nodes with the following properties:

- E increases when the shape is deformed away from its equilibrium state;

- E decreases when the shape is deformed closer to its equilibrium state;
- E has a global minimum when the cloth is in the equilibrium state;

Starting from an initial (rectangular) shape, the nodes were "relaxed" in sweeps, i.e. every node was moved so that the overall energy E decreased. Feynman was the first to address the problem of collision detection for systems including deformable bodies (in his example cloth) and non-rigid obstacles. He used simple manual adjustments of node positions when they were detected to have penetrated the obstacles.

First rigorous theoretical treatment by Terzopoulos et al: 1987

Terzopoulos et al [7] published the first fundamental work in the field of modelling deformable bodies for computer graphics in 1987. The theory of elasticity was used to derive a general simulation system of differential equations, capable of modelling the dynamics of one-dimensional (curves), two-dimensional (such as cloth) and three-dimensional (solids) bodies under the action of applied forces. The Lagrange form of motion equations was used:

$$\frac{\partial}{\partial t}\left(\mu\frac{\partial r}{\partial t}\right)+\gamma\frac{\partial r}{\partial t}+\frac{\delta\varepsilon(r)}{\delta r}=f(r,t) \qquad (1.1)$$

Here the first term on the left is the inertial force due to the distributed mass, the second term is the damping force due to dissipation, the third term is the elastic force due to the deformation of the body from its original (undeformed) shape, the term on the right is the resultant of the external forces acting at the point. The variational derivative of the potential energy of the deformation $\varepsilon(r)$ is given by $\frac{\delta\varepsilon(r)}{\delta r}$. It is the force due to the deformation of the object.

$\varepsilon(r)$, as the measure of the deformation, was expressed via the measure of the deformation of the metric tensor for solids, the deformation of the both metric and curvature tensor for surfaces and the deformation of the metric, curvature and torsion tensors (scalars) for curves.

The simplified expression for the variational derivative was used and a semi-implicit integration procedure was suggested to solve the system of equations numerically. Collisions with external objects were dealt with by introducing the potential field around the

objects that produced repulsive forces with their magnitude growing exponentially as the objects approached each other.

This work was the first theoretically rigorous treatment of the subject and it is probably the mostly frequently quoted paper on elastic deformations for computer graphics. The proposed approach was improved and extended the following year by Terzopoulos et al [12]. Rigid body translations and rotations were included in the system of equations, and treatment of other phenomena such as plasticity, viscoelasticity and fracture was investigated.

Another theoretical approach to modelling cloth by Aono: 1990

In 1990 Aono published his research findings in the modelling of cloth [18]. His work was another theoretically rigorous treatment of the way the simulation system (the system of equations to describe the evolution of the system in time) was produced. Aono's approach was rather different from Terzopoulos [7], in that Aono used the classical constituent equations (strain-stress relationships) from the theory of elasticity:

$$\begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{xy} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \end{pmatrix} = \begin{pmatrix} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & \frac{1}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2G} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2G} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2G} \end{pmatrix} \begin{pmatrix} \tau_{xx} \\ \tau_{yy} \\ \tau_{zz} \\ \tau_{xy} \\ \tau_{yz} \\ \tau_{zx} \end{pmatrix} \qquad (1.2)$$

Here E is the Young's modulus of elasticity in tension, G is the modulus of rigidity (or the modulus of elasticity in shear) and ν is the Poisson ratio.

A number of simplifications were used, namely:

- The cloth was homogeneous (that is, the same material properties throughout the cloth), isotropic (the material properties were the same in any direction) and linearly elastic in its initial shape.

- The cloth was in equilibrium at any time under given applied and inertial forces (D'Alembert's principle).

- The cloth was a perfectly thin surface and never expanded or contracted along its surface normal.

Later on, however, it was shown how the model could be modified to deal with real life situations of inhomogeneity, anisotropy and viscoelasticity, that is, with the materials which may have different material properties in different parts, in different directions in space and which when deformed exhibit more complex behaviour.

Aono's approach was novel to the previous ones in that he used the formulation where the mechanical parameters of cloth were present explicitly. His model was also capable of dealing with the situations of inhomogeneity, anisotropy and viscoelasticity, which earlier work could not model.

Emergence of several distinct areas of research in modelling flexible bodies

It became clear by the end of 1980s that the whole problem of developing an efficient physically based systems for modelling flexible bodies could be subdivided into a number of relatively independent areas:

- The simulation system itself. This mainly consists of the mathematical model of the simulation system (including the mathematical description of internal and external factors such as forces, material properties and constraints) and a set of mathematical equations (the equations of motion) describing the evolution of the system over time.

- A set of numerical integration methods for the practical solution of the equations of motion. The numerical methods are distinguished by their stability, speed and precision.

- Collision detection and response. This area includes self-collisions, collisions between objects in the simulation system and ways of dealing with these collisions (generally described as collision response).

- The problem of control of the simulation system, that is the ability by the animator to specify the desired evolution in a more natural and accessible form (for example, by giving the starting/ending point, by specifying certain criteria the system has to

minimise throughout the motion), rather than by the means of adjusting the numerous simulation parameters and forces.

Simulation system for flexible bodies based on global deformations of the rest shape: 1992

Witkin and Baraff [27] developed a method of describing the deforming bodies in terms of the global deformations of the rest shape, based on earlier work by Welch and Witkin [19]. In Witkin and Baraff's method the changes in the body's shape were represented by global parameterised geometric approximation.

The method proposed by Witkin and Baraff was positioned between the extreme cases of rigid body motion simulation and "local" nodal simulation models. As the result, the authors claimed to avoid the stiffness problems encountered in other models based on nodal approximations, while still allowing a rich variety of deformations to be modelled efficiently.

A method to handle collisions based on an analytical contact forces developed by Baraff was also proposed.

Particle based method by Breen et al: 1994

In 1994 Breen, House and Wozny [38] produced a distinctively different non-continuum model for predicting the drape of cloth. They treated the cloth as a collection of interacting particles with the approach being based on the fact that fundamentally cloth is made of an interlocking network of yarns and threads. The intersections of the yarns and threads could then be treated as heavy particles with the friction forces between the threads being modelled by the forces of interaction between the particles.

The energy function (as the measure of deformation and which is related to forces between particles) was calculated for every particle:

$$U_i = U_{repel_i} + U_{stretch_i} + U_{bend_i} + U_{trellis_i} + U_{gravity_i} \qquad (1.3)$$

The energy function was then minimized to produce the equilibrium shape of the modelled cloth.

The authors made a specific effort to produce a system capable of accurate modelling of real clothes (cotton, wool). They developed a way to "tune" the energy function for particular

cloths, based on utilisation of mechanical data produced by the Kawabata Evaluation System [1].

Below the specifics of their approach are outlined:

- It is straightforward to implement.
- It only deals with static modelling. However, a straightforward extension to include dynamics is possible ([58]).
- It describes the means to incorporate the results of tests produced by the Kawabata Evaluation System, thus allowing them to capture the specific draping behaviour of real materials (cotton, wool, silk etc).

A consistent and efficient cloth model by Baraff and Witkin: 1998

Baraff and Witkin [72] developed a technique for modelling cloth, which accumulated the latest achievements to date. It could be regarded as a fundamental work on the subject. The authors proposed:

- Efficient and simple formulation of internal forces through the internal energy.
- Consistent and efficient damping forces.
- Implicit adaptive step integration technique (modified backward Euler plus conjugate gradient method) for stability and speed (large steps could be taken).
- Consistent treatment for self-collisions (coherency-based bounding box approach) and collisions with immobile objects (through stiff springs, which is not a problem because implicit integration can handle stiff springs easily).
- Efficient treatment of constraints that are enforced exactly (at all times).
- Consistent treatment of friction.
- Efficient surface meshing through triangulation.

1.2.3 Brief summary of major developments

After the pioneering work by Weil [5] and Feynman [6] on cloth modelling, the first rigorous treatment of the subject, based on the classical theory of elasticity, was proposed by Terzopoulos et al [7] in 1987. Models based on Lagrangian mechanics were proposed for

all practical cases: 1D curves, 2D surfaces and 3D solids. The method was later extended by Terzopoulos et al [12] to include a framework for the treatment of inelastic deformations (viscoelasticity, plasticity and fracture). Additionally, the reference frame was allowed to undergo rigid body movement along with the body (thus widening the range of phenomena that can be modelled).

At the same time, another approach, based on presenting the object to be modelled as a collection of mass points with springs between them, started to become a popular method of modelling flexible bodies. Miller was one of the first to apply it to model the dynamics of snakes and worms [9] in 1988.

The mass-spring approach has a number of distinctive features:

- It is quite simple;
- The limitation of the continuum approach of small deformations away from the reference shape was removed (modelling snakes was a good example to demonstrate this);
- It can be adapted for a variety of situations: from modelling snake and worms [9] to modelling motion of artificial fishes [41] and cloth [38, 39, 53, 58]

Thalmann and colleagues started their research in modelling cloth on synthetic actors in 1991 [20]. They initially used the (slightly modified) approach proposed by Terzopoulos [7,12] but later abandoned it [46] in favour of the spring-mass approach. They developed a self-collision detection and response algorithm [20], which is particularly important in the case of actors wearing cloth that can overlap multiple times. The algorithm was later significantly improved [43]. An impressive system for modelling cloth on synthetic actors was developed as a result throughout the years [73], with such features as ability to design cloth patterns in a 2D environment (with cutting and seaming capabilities) and then import them into a 3D world and put it on a actor interactively, and sophisticated self-collision and cloth-to-body collision handling.

Aono [18] in 1990 used classical theory of elasticity to develop a model to simulate wrinkles and their propagation in cloth. The model was only suitable for modelling globally small deformations.

In 1992 Witkin and Baraff [27] proposed a method for describing the deformable bodies (solids) in terms on the global deformations of the rest shape, when the changes in the body's shape were represented by global parameterised geometric approximation.

Breen et al [38, 39] in 1994 developed a particle-based method (mass-spring model) for modelling draping of cloth. The proposed method was based on energy minimisation and therefore was very slow.

In 1995 Provot [53] employed a mass-spring method to model cloth. Inverse kinematics was used to improve visual realism (to resolve the problem of super-elasticity) and account for non-linear stretching behaviour of real cloth.

In 1996 Eberhardt et al [58] used the same particle-based approach as Breen et al [38,39] to model draping of cloth. The model was improved to include dynamics and fast numerical solution of the equations was proposed so that interactive animation rates could be achieved.

In 1999 James and Pai [79] proposed a method for modelling deformable objects (solids) based on boundary integral formulation. This method is a novel approach to the problem, it seems to be best adopted to the situations when deformable object is a sort of inflated soft solid being manipulated by a localised external force (like pulling).

Recent articles by Witkin and Baraff [72] and Desbrun et al [75] can be regarded as the current state-of-the-art in cloth modelling. Fast (interactive rates), efficient and stable models were proposed for animating cloth with advanced self-collision detection and response.

Generally, it seems the research community has come to the conclusion that theoretical foundations for fast realistic modelling of cloth have been well established and nothing new can be added. A fast and stable method for cloth animation is some sort of mass-spring model (for speed) combined with implicit integration technique (for stability and large integration steps).. It is not even that important how the collisions are handled: stiff repulsive springs, penalty methods, inverse kinematics; all these can be stably manipulated by implicit integration and therefore which collision handling technique to choose becomes a matter of personal preference and ease of use in a particular applied problem.

It is not as simple in case of flexible solid modelling. Three basic methods were proposed:

- Terzopoulos et al [7] used Lagrange dynamics formulation from the classical theory of elasticity for the case of 3D solids. The method does not explicitly include standard elastic parameters ν, E and does not provide any description on how to deal with the boundary forces.

- Witkin and Baraff [27] approximated the possible deformations by global parameterised geometric approximation. The method's formulation placed an obvious restriction on the class of phenomena it is capable of adequate modelling.

- James and Pai [79] flexible body deformation representation through boundary integral formulation. It also seems to be restricted to specific class of objects ("inflated" etc) and is best suited for specific simulations (when volume is preserved during the deformation and so on).

However, in spite of the volume of these research efforts, problems still exist. Whilst in some areas strong results have been achieved (for example, cloth simulation on synthetic actors by the Thalmann's group), some other areas of research received little attention. Modelling flexible 3D solids is an example of such an area. Generic theoretical treatments of 3D flexible solid modelling have been suggested [7,12,27] but very few practical implementations have been undertaken.

One of the main reasons for this is the fact that the suggested approaches contain serious limitations. Terzopoulos's method [7,12] consists in fact of two formulations: the primary and the hybrid one and has several limitations, such as the use of large number of parameters that need adjusting on a trial and error basis and the lack of description of how to deal with boundary forces and internal energy dissipation. The global geometric deformation approach suggested by Witkin and Baraff [27] is obviously limited to dealing with systems where the types of deformations are only those presentable by the employed geometric parameterisation. In their work the authors themselves only considered first and second order polynomial deformations.

1.3 The Structure of the Thesis

The thesis consists of five chapters.

It begins with the introduction in Chapter 1. We give a general overview of applied computer modelling. This is then followed by a historical overview of the major developments in the area of the applied modelling over the last twenty years in section 1.2.

Since our theory of FlexyMatter is an extension to the classical theory of elasticity, in Chapter 2 we give a complete presentation of the classical theory of elasticity in the form that is most suitable for this thesis. We begin with an overview of modelling three-dimensional solids in general in section 2.2. We point out major issues involved and give a general definition of the modelling problem. We then proceed to outline a mechanical definition of the modelling task in section 2.3.

In the next section 2.4 we give a full mathematical definition of the modelling task. We give a mathematical definition for all the quantities involved in the modelling task: time, the solid itself and the global coordinate system. We then introduce and discuss the mathematical equivalents to the physical quantities involved (such as, the mass or density, the velocity, the force and so on).

In section 2.5 we provide a simple derivation of the notion of the elastic force. We do it via the basic law of motion for our solid based on Newton's second law of motion. We also point out at the end of the section that even though we are able to establish the existence of the elastic force and several of its main properties very quickly and simply, we are still some way away from being able to derive an expression to allow us to calculate it for any solid.

In the section 2.6 we formally introduce the notion of a curvilinear coordinate system and all the mathematical apparatus associated with it: covariant and contravariant base vectors and coordinates, coordinate transformations, tensors, covariant derivative, divergence of a vector field, integration and the Gauss' divergence theorem.

By this stage we have all the required mathematical apparatus and in section 2.7 we proceed to introduce the basic entities from the classical theory of elasticity required to derive the expression for the internal elastic force. These include displacement field, strain and stress tensor. We conclude the section with a presentation of the constitutive equations binding the stress and strain tensors and relating to the material properties of a solid.

In section 2.8 we introduce the fundamental equations of the classical theory of elasticity. We begin with the equations of motion for the internal points of any solid. We then follow with the so-called boundary conditions – the equations describing the equilibrium of forces on the boundary of the solid. The section is concluded with a full set of equations for all points of the solid.

In the next secton 2.9 we provide a simple overview of the limitations of the classical theory of elasticty, including the limitations on the displacement gradients and its inability to model the rigid body motion and rotation.

The next Chapter 3 is dedicated to an overview of the existing practically established modelling techniques, developed by the research community over the last twenty years. We first divide all the methods into two groups: the local modelling methods (section 3.1) and the global modelling methods (section 3.2). The methods are distinguished by the presence of the global terms, such as, the centre of mass and the inertia tensor, in their equations of motion: those methods that have such global terms are referred to as the global group and those that do not – as the local group. For each group we provide a detailed overview of the most illustrative practical methods in the group. The last section 3.3 is a conclusion; here we outline the problems that still exist in the practical modelling methods and point out that our theory of FlexyMatter is able to successfully resolve most of those problems.

Chapter 4 provides our argument for the need for a new modelling theory. We begin with a complete and detailed analysis of the limitations of the linear classical theory of elasticity in section 4.2. Based on this analysis we propose and discuss one possible extensions to the classical theory that address some or all of the outlined limitations: the global alignment approaches in section 4.3. At the end of the section we point out that this global alignment approach is not an adequate solution; it does not solve all the problems of the classical theory and introduces some artefacts of its own.

Chapter 5 presents formally the theory of FlexyMatter as an extension to the classical theory of elasticity, that is, fully based on the terminology and ideas introduced in chapter 2. We show that the local alignment approach, which can be considered the basis of the theory of FlexyMatter, is an extension to the linear classical theory that removes all the problems with the linear classical theory outlined in section 4.2.

The next chapter 6 provides a full description of a practical implementation of our modelling technique: from the discretization of the modelling problem in section 6.1 to the practical implementation of the numerical solution in section 6.3.

In section 6.4 we present a set of results, obtained using the theory of FlexyMatter. The section shows our theory in action, providing a practical proof of the validity of the concept.

Chapter 7 contains the summary of the results achieved in this thesis. It outlines the areas where the theory of FlexyMatter has advantages over the existing modelling techniques, and provides an overview of future work.

Appendix A contains an overview of numerical methods which may be of use to any one wishing to develop their own modelling system based on our theory of FlexyMatter.

In the Appendix B we provided a comprehensive overview of the theory of Hyper-Matter by Smith and Paddon.

2. Classical Theory of Elasticity

2.1 Introduction

Any modelling is based on prediction. Knowing the initial state, the applied forces and laws to which the system will be subjected to during its future evolution (change with time), the task is to predict the state of the system at any given time after the commencement of the experiment.

Specifically, in the case of modelling 3D elastic solids given the initial position, shape and velocity of the object, the forces acting on it at all times together with the laws that will govern the deformation, the task is to predict what the object will look like after the start of the experiment.

In our real life, we are very used to seeing objects and their evolution under applied forces. In fact, in many cases we can quite accurately predict beforehand what will happen to an object if we applied given force(s). We know that if we drop a rubber ball on the floor, it will bounce back to a height that depends on the properties of the rubber and the floor it hits. We know if we hit it very hard, it may break and its bouncing (or elastic) properties will dramatically change. We know that a wooden beam constrained at its ends will bend downwards in the middle if we put a weight on it. We can even calculate the bend as a function of the applied weight and the thickness of the beam.

From these examples, we can draw the following conclusions:

- The behaviour of the objects that we deal with in real life is predictable. If we can derive laws which the objects obey as they evolve subject to the action of given forces (regardless of the way the laws were obtained), we can use those laws to predict the evolution of the objects at any later point in time. In other words, we

assume the validity of the following assumption: there exists a definite law for our physical system that governs its evolution subject to given external forces.

- It is possible to derive a law based on practical experience or experiments for a certain linited range of objects and/or situations, and then extend it to a wider range.

These are important assumptions: the notion of predictability (or determinism) must be treated with care. A number of theories that appeared during the last hundred years or so (namely quantum mechanics, theory of chaos), show that in some cases a deterministic law does not exist. For example, one cannot predict the position of an electron at any moment in time with an absolute certainty; one can only say where it is most likely to be, even after its exact position has been measured by an experiment.

Very often the law that has been derived for a certain range of situations can only be extended into a relatively wider range of situations, but it cannot be extended without limit. We can, for example, determine experimentally the law of addition of velocities (if two objects are flying in one line with the velocities of v_1 and v_2 respectively then the velocity of second ball as seen from the first ball will be $v_2 - v_1$). But, according to the general relativity theory, it is only an approximation, and it can only be used when both velocities are small comparing to the velocity of light. However, it is a very good approximation for small velocities, and as long as we are aware of the range of its applicability, there is no need to use the general one as given in the general relativity theory.

2.2 Modelling 3D elastic solids: historical overview

The classical theory of elasticity (as well as most other theories of modelling the evolution of systems under the action of applied forces) is based on Newton's second law, which states that for any material object its mass multiplied by its acceleration is equal to the resultant of the forces applied to it. This law was established by Newton in the 18^{th} century and has been used as a fundamental law of motion ever since.

The classical theory of elasticity was mostly developed in 19^{th} century and until the beginning of 20^{th} century was believed to be applicable to and valid for a wide range of systems: from very small to very large (because it is based on Newton's second law which was assumed to be universally valid). However, the quantum theory and the general theory of relativity set the limits of its applicability. In fact, it was demonstrated that it is really a

special case of a more general theory (the general theory of relativity) and is not applicable at all on a small (atomic) scale.

However, the theory works well as long as the objects are not too small (significantly larger than an individual atom), not too large (significantly smaller than a star or a planet), and do not move too fast (significantly slower than the speed of light). And because it is considerably simpler than the general theory of relativity, it is still a preferred choice in applied modelling.

2.3 The modelling task: mechanical formulation

We have already presented a complete mechanical formulation of our modelling problem. This section is a quick reminder.

Given a solid, its distribution of mass, material properties, initial position and velocity, the distribution of volume and boundary forces (given at all times, not just at the start of the experiment), the task is to determine the position, velocity and the deformation of the solid throughout the experiment.

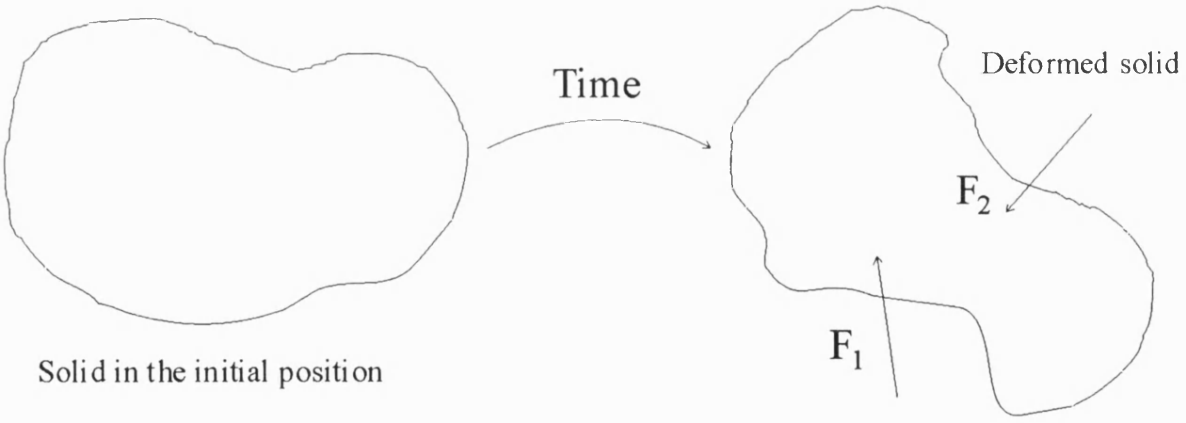


Figure 2.1: Solid deforming with time under the action of the applied forces $\mathbf{F}_1$ and $\mathbf{F}_2$

Obviously, solids have a number of material properties, including: elasticity (resistance to stretching and squashing directly proportional to the actual deformation), viscosity (material resistance proportional to the rate of deformation), plasticity (internal structural changes that result in the solid assuming new shape). All these properties will need to be taken into account during the modelling process.

2.4 Mathematical Model

The theory of elasticity is a mathematical theory. It is applied to a mathematical equivalent of our problem and operates in terms of mathematical quantities: material points and their coordinates, density of the material at a point, volume and surface forces acting on a point and so on. We therefore need to start by converting the previously described mechanical formulation of the problem into a mathematical form.

2.4.1 Time

Our modelling problem is dynamic: the system changes with time. We should therefore describe how time would be accounted for in the mathematical model.

We will denote time by symbol t and assume, without the loss of generality, that the experiment starts at $t = 0$. Time therefore will be a parameter in our experiment (a real number t, such that $t \geq 0$) and all the quantities and physical entities (position, forces etc) will be functions of this parameter. We are not concerned here with the units of measurement of t; they could be seconds, minutes or days. All we need to know is that $t = 0$ is the start of the experiment and $t > 0$ is a certain point into the experiment. Also, if we have $t_1 \geq 0$, $t_2 \geq 0$ and $t_2 > t_1$, we can definitely say that t_2 happened after t_1.

2.4.2 Mathematical model of the body

It is well known that physically any solid consists of a very large number of very small elements (atoms), bound together by the internal forces. To represent the body mathematically, we will need the following assumption:

- We assume it is continuous: that is, it can be indefinitely divided into pieces and inside each piece there will always be material of which the body is made of. This is clearly an idealisation of the real situation: a real physical body cannot be subdivided indefinitely. As soon as one reaches atomic level, the body starts to look more like a lattice, rather than a smooth substance. However, in order to formulate the theory we need this assumption, since we will be performing mathematical operations (taking derivatives and integrals) that require this notion of continuity.

In the view of this assumption, we will represent a body as a continuous region in space, with all the points of the region belonging to the body and denote it as B.

2.4.3 Worlc coordinate system

We will now ntroduce a global (world) coordinate system into our world where we will be modelling our body.

We will assume we have a global rectangular Cartesian coordinate system K, which will cover the whcle region of space where we conduct our experiment. The starting point for the coordinate system can be chosen arbitrarily but it is fixed (see Figure 2.2).

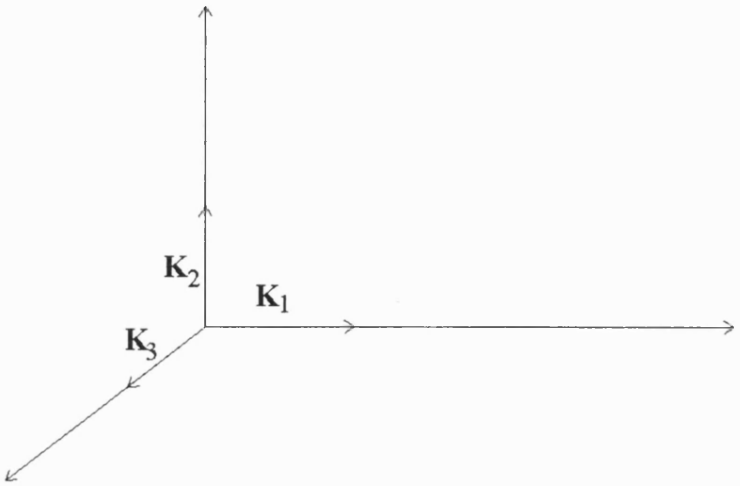


Figure 2.2: Global Cartesian coordinate system

$\mathbf{k}_1$, $\mathbf{k}_2$ and $\mathbf{k}_3$ are the base vectors. K is Cartesian, therefore:

$$\begin{aligned} |\mathbf{k}_i| &= 1, \forall i = 1,2,3 \\ \mathbf{k}_i \times \mathbf{k}_j &= \delta_{ij}, \forall i, j = 1,2,3 \end{aligned} \tag{2.1}$$

That is, they are all of unit length and perpendicular to each other. δ_{ij} here is the Cronecker delta, for which the following is satisfied:

$$\delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \forall i, j = 1,2,3 \tag{2.2}$$

2.4.4 Important note on notation

Later in the chapter, we will be dealing with many quantities, which are scalars, vectors or tensors. We will use the following notation.

- Vectors will be displayed in bold: $\mathbf{v}$, $\mathbf{k}_i$ etc with the understanding that the vectors are in the global coordinate system K and are defined (that is, we know how to find the oriented arrow that they represent).

- Scalars will be displayed in italic: a, a_i, δ_{ij}, δ^i_j etc for any fixed values of indices presented by Latin letters (i and j in these examples).

- Any vector **a** in any given coordinate system can be represented by its coordinates. In other words, it can be presented in the following form:

$$\mathbf{a} = a^1\mathbf{k}_1 + a^2\mathbf{k}_2 + a^3\mathbf{k}_3 \qquad (2.3)$$

  In this case, we can refer to the same object (vector) by **a**, or by its coordinates a^i. Therefore, the other notation that will be used extensively is indexed quantities like a^i which will refer to the whole object rather than an individual coordinate a^i when i is fixed.

Our body B is positioned somewhere in the space:

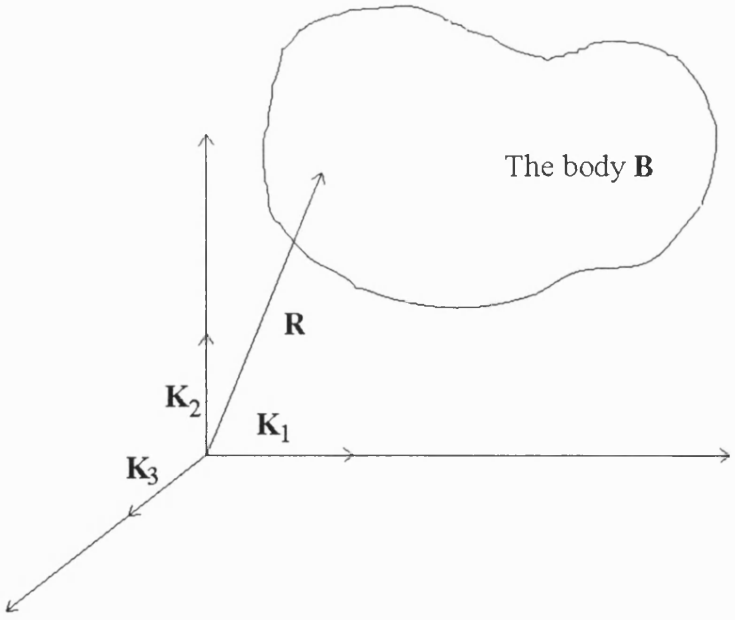


Figure 2.3: Body B in our global coordinate system

As it was mentioned before, our body B is a mathematical model of a physical body. B is continuous and encompasses all points in our space belonging to the region of space occupied by the physical body.

In Figure 2.3, **R** is a position vector of a point belonging to B : $\mathbf{R} \in$ B

2.4.5 Mathematical equivalents of physical quantities that apply to solids

There are at least two types (later we will see there are more) of mechanical quantities that apply to solids (or parts of solids):

- The first one consists of scalars. A typical example of a scalar is the material density ρ. Physically, it is measured in unit mass per unit volume, and describes the concentration of the material within the solid. In our continuous model, we define ρ as a function of a point $\mathbf{R} \in \mathrm{B}$:

$$\rho(\mathbf{R}) = \lim_{dV \to 0} \frac{m(dV)}{dV} \tag{2.4}$$

  dV here is the volume of any section of the body that contains **R**, and $m(dV)$ is the mass of this section. One can see that this definition is again an idealisation of a real physical situation; as for real solids, one cannot continually shrink the section around a point in the body without eventually running out of the material of which the body is comprised.

  The connection between this idealised continuous density and the real mass is quite simple, we can measure this by an experiment. Let us suppose we have $\rho(\mathbf{R})$ for a point $\mathbf{R} \in \mathrm{B}$ and a small section of material around the point. Then the real mass will be:

$$m \cong \rho(\mathbf{R}) dV \tag{2.5}$$

  where dV is the volume of the section.

  Another typical example of a scalar quantity defined for a solid is temperature T. We will assume the temperature T is a scalar function of a vector argument:

$$T = T(\mathbf{R}), \mathbf{R} \in \mathrm{B} \tag{2.6}$$

  That is, it is defined for every point $\mathbf{R} \in \mathrm{B}$. We will not deal with temperature in this thesis, and therefore it is only given here for illustrative purposes.

- The second type consists of vectors. The most typical example of this type is a force **F**. We suppose that any force is a vector function of a vector argument:

$$\mathbf{F} = \mathbf{F}(\mathbf{R}) = F^1(\mathbf{R})\mathbf{k}_1 + F^2(\mathbf{R})\mathbf{k}_2 + F^3(\mathbf{R})\mathbf{k}_3, \mathbf{R} \in \mathrm{B} \tag{2.7}$$

Which means it is a vector defined for every point $\mathbf{R} \in \mathrm{B}$. Again, the real force that can be measured experimentally would act on a piece of material, not a point, but the connection between the two is simple. We take a small section containing a point $\mathbf{R} \in \mathrm{B}$, and then the real force is:

$$\begin{aligned} \mathbf{F} &\cong \mathbf{F}(\mathbf{R})dV \\ \mathbf{F} &\cong \mathbf{F}(\mathbf{R})dS \end{aligned} \tag{2.8}$$

where dV is the volume of the section and dS is the area of the surface of the section. The first relationship is for volume forces (the forces that "act" on the volume and are measured in force units per unit volume, for example gravity) and the second one is for surface forces (the forces that "act" on the surface and are measured in force units per unit area, for example, the surface friction force).

Another example of a vector quantity is velocity $\mathbf{V}(t)$ of a point at a time instant t:

$$\mathbf{V}(t) = \mathbf{V}(\mathbf{R}(t)) = \frac{d}{dt}\mathbf{R}(t) \tag{2.9}$$

2.4.6 Functional mapping for the solid

In our model, a body B is a set of points **R**. So far we made no assumptions of what sort of set it is. But, once we have departed from the physical world into our idealised world, we need to provide a mathematical definition of what sort of set will constitute a physically realistic body. The following examples illustrate the need for a formal definition:

a) Can a single point or a set of single points be a valid body B ?

b) Can a line or a two-dimensional curve be a valid body B ?

We will now derive a formal definition of a three-dimensional solid: a solid that is not infinitely thin at any of its points. This is really a clarification of the idea of a continuous three-dimensional solid.

First we will define the notion of a three-dimensional parallelepiped – a simple object, which we will later use as domain in the definition of the functional mapping for our solid.

Definition (2.1)

The following is a definition of a two and three-dimensional parallelepiped (called domain parallelepipeds):

$$I = I[A^1, A^2, A^3] = I^3[A^1, A^2, A^3] = \left\{(x, y, z) : x \in [0, A^1], y \in [0, A^2], z \in [0, A^3]\right\}$$

$$\begin{aligned} & I^2[A^1, A^2] = \left\{(x, y) : x \in [0, A^1], y \in [0, A^2]\right\} \\ & A^1, A^2, A^3 \in \mathrm{R} \\ & A^i > 0 \end{aligned} \tag{2.10}$$

When I is used instead of the full $I[A^1, A^2, A^3]$, we assume that certain definite values $A^i : A^i > 0$ are chosen and fixed.

We can now define a three-dimensional solid.

Definition (2.2)

A collection of points B is called a valid three-dimensional solid, if and only if there exists a one to one mapping **R** such that:

$$\begin{aligned} & \mathbf{R} = \mathbf{R}(x^i) : I[A^1, A^2, A^3] \to \mathrm{B} \\ & \mathbf{R} \in C_I^1 \end{aligned} \tag{2.11}$$

R maps $I[A^1, A^2, A^3]$ onto our three dimensional body B (for certain values of A^i), it is continuous and differentiable with a continuous derivative (the derivative on the boundary is defined as a limit of a derivative at any internal point approaching the boundary point along any path: a curve, starting at the internal point and ending at the boundary point). Treatment of boundaries in the theory of elasticity is a separate problem and it will be dealt with in section 2.8.2.

Here, x^i is used to denote a vector $\mathbf{x} = x^1\mathbf{k}_1 + x^2\mathbf{k}_2 + x^3\mathbf{k}_3$ as a whole, as opposed to each individual component.

In the light of this definition, one can say that the answer to the above questions (a) and (b) is no, they are not valid solids.

In this thesis, we will only deal with valid solids according to (2.2). This definition is, in fact, slightly restrictive. For example, the set $S = I \cap \{P\}$, where P is a point from the

boundary, that is, the domain parallelepiped without one point on its boundary, will not be a valid solid according to this definition. However, the aim of this thesis is to demonstrate the fundamental ideas of our approach rather then provide a generic theory of modelling solids of any kind. And therefore for simplicity we will use this definition.

One important note has to be made here. If we have a valid body and, therefore, there exists a mapping **R** satisfying the conditions from the definition, one might ask a question: is this **R** unique?

It can be demonstrated, that, in fact, if there exists one mapping **R**, there exists an infinite number of different mappings. Indeed, it is not difficult to show that there exists an infinite number of mappings $\mathbf{R} = \mathbf{R}(x^i) : I \to I, \mathbf{R} \in C_I^1$. Therefore, if there exists a $\mathbf{R} = \mathbf{R}(x^i) : I \to B, \mathbf{R} \in C_I^1$, one can superpose it with any $\mathbf{R} = \mathbf{R}(x^i) : I \to I, \mathbf{R} \in C_I^1$.

We can also note that when we say that we have a mapping function **R**, we imply that we have a set of A^i specified as well (not necessarily precisely). For example, when we say we have a mapping **R** covering a region in space, A^i do not need to be defined precisely. All we can say about them is: "they are such that our region is covered". However, when we say that we have a mapping for a specific body B, it would imply the existence of a mapping function **R** as well as a set of A^i defining its domain.

2.4.7 The full mathematical statement of the modelling problem

We can now state the mathematical equivalent of the modelling problem given in section 2.3, using all the new definitions and considerations we have introduced so far.

At the initial time $t = 0$, we have the following:

- A body B, given by a mapping function **R**. As shown above, there can be many possible mappings **R**, we choose one and fix it.

$$\mathbf{R} = \mathbf{R}(x^i) : I[A^1, A^2, A^3] \to \mathrm{B} \tag{2.12}$$

- A volume force field **F** (that is, the force defined for every internal point) at any time t:

$$\begin{aligned} &\mathbf{F} = \mathbf{F}(\mathbf{R}, t), \forall \mathbf{R} \in \mathrm{B}, t \geq 0 \\ &\mathbf{F} = \mathbf{F}(\mathbf{R}(x^i), t) = \mathbf{F}(x^i, t), \forall x^i \in I, t \geq 0 \end{aligned} \tag{2.13}$$

These two definitions of the force field are equivalent: if force is given for every position vector $\mathbf{R}$, and $\mathbf{R}(\ldots)$ is a mapping and, therefore, $\mathbf{R} = \mathbf{R}(x^i)$, the force $\mathbf{F}$ is also a function of x^i (subject to the choice of $\mathbf{R}$, of course).

- A surface force field $\mathbf{S}$ (that is, the force defined for every boundary point) at any time t:

$$\begin{aligned} &\mathbf{S} = \mathbf{S}(\mathbf{R},t), \forall \mathbf{R} \in \partial \mathrm{B}, t \geq 0 \\ &\mathbf{S} = \mathbf{S}(\mathbf{R}(x^i),t) = \mathbf{S}(x^i,t), \forall x^i \in \partial I, t \geq 0 \end{aligned} \tag{2.14}$$

where $\partial \mathrm{B}$ is the boundary of the body, and ∂I is the boundary of the domain parallelepiped B.

- The material density ρ at any time:

$$\rho = \rho(\mathbf{R},t) = \rho(\mathbf{R}(x^i),t) = \rho(x^i,t), \forall t \geq 0, \forall x^i \in I \tag{2.15}$$

- The initial velocity $\mathbf{V}$:

$$\mathbf{V} = \mathbf{V}(\mathbf{R}) = \mathbf{V}(\mathbf{R}(x^i)) = \mathbf{V}(x^i), \forall x^i \in I \tag{2.16}$$

The problem can then be stated formally as follows:

Given the body B, mapping function $\mathbf{R}(x^i), \forall x^i \in I$, volume force field $\mathbf{F}(x^i,t), t \geq 0$, boundary force field $\mathbf{S}(x^i,t), t \geq 0$ and the velocity field $\mathbf{V}(x^i), \forall x^i \in I$, find $\mathbf{R}(x^i,t), \forall x^i \in I, t > 0$:

$$\mathbf{R}(x^i,\circ) \in C^1_I \wedge \mathbf{R}(\circ,t) \in C^2_{[0,\infty[} \tag{2.17}$$

That is, find a vector function $\mathbf{R}(x^i,t)$, which is once differentiable by $x^i, i = 1,2,3$ and twice differentiable by time t with continuous derivatives.

This function $\mathbf{R}(x^i,t)$ must for every instant $t > 0$ represent the deforming body as it deforms under the action of applied forces $\mathbf{F}(x^i,t)$.

For every instant in time t we have a different (deformed) body. It therefore is a function of t and we will denote it $\mathrm{B}(t)$. The following Figure 2.4 provides a graphical illustration:

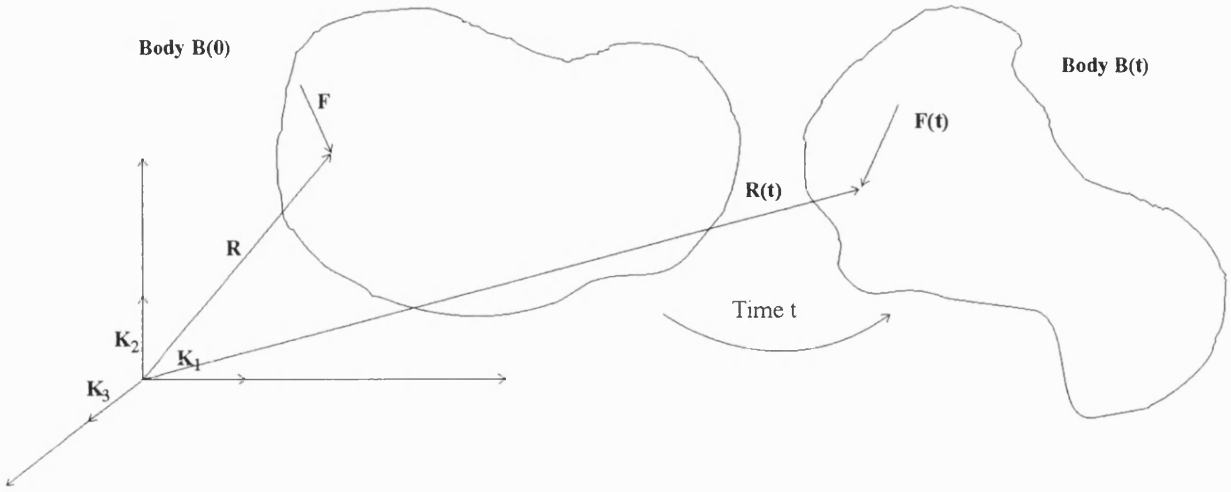


Figure 2.4: Mathematical statement of modelling problem

Note that we can formally write:

$$\mathbf{R}(x^i) = \mathbf{R}(x^i, 0) \tag{2.18}$$

Similarly we will have:

$$\mathbf{V}(x^i, t) = \frac{d}{dt}\mathbf{R}(x^i, t), t \geq 0 \tag{2.19}$$

$\mathbf{V}(x^i, t)$ is the velocity of the body (all of its points) at the instant of time t. We do not need to find it explicitly as it depends on $\mathbf{R}(x^i, t)$.

2.5 Preliminary Analysis and Simple Derivation of Elastic Force

2.5.1 The basic law of motion

The primary purpose of this section is to introduce the law of motion for a solid, and introduce the notion of an internal force. We will also give general considerations about the internal force before the introduction of the significantly more complex part of the elasticity theory, dealing with curvilinear coordinate systems, tensors and so on.

It is possible to use simple considerations and develop some understanding of the nature and the form of the law that governs the evolution of the solid through the time.

For simplicity, we will assume the body evolution is such that its boundary remains stationary. That is, there is a constraint that the boundary does not move. The reason we need to make this assumption is simple: as we will show later, handling the boundary adds

complexity that will obscure the main purpose of this section, which is to do a simple analysis to provide an understanding of the law of motion for the solid.

Our primary tool here is Newton's second law of motion: the total force acting on each small element will be equal to the mass of the element multiplied by its acceleration.

Let us choose such a small volume element of the body Ω and assume that it is purely internal (that is, it does not have common points with the solid's boundary), dV will be its volume. We can now write down the following equation:

$$\mathbf{a}\,dm = \mathbf{\Phi} dV \tag{2.20}$$

where $\mathbf{a}$ is the average acceleration of the element, dm is its mass and $\mathbf{\Phi}$ is the total average force acting on the element.

Average acceleration of the element is:

$$\mathbf{a} = \frac{1}{dV}\int_{\Omega}\mathbf{a}(x^i,t)\Delta V \tag{2.21}$$

We assume our volume element is small, thus, the acceleration can be assumed to be almost constant and, therefore, we can choose a point inside the element (for example the centre of mass $\mathbf{R}_c = \mathbf{R}(x_c^i)$) and say that the acceleration throughout the whole of the element is equal to that of the centre of mass $\mathbf{R}_c$. We can therefore write:

$$\mathbf{a} \cong \frac{1}{dV}\int_{\Omega}\ddot{\mathbf{R}}_c\Delta V = \ddot{\mathbf{R}}_c \tag{2.22}$$

Here dot means a derivative by time as usual.

Let us now see what this total force $\mathbf{\Phi}$ consists of by considering the balance of forces acting on the volume element.

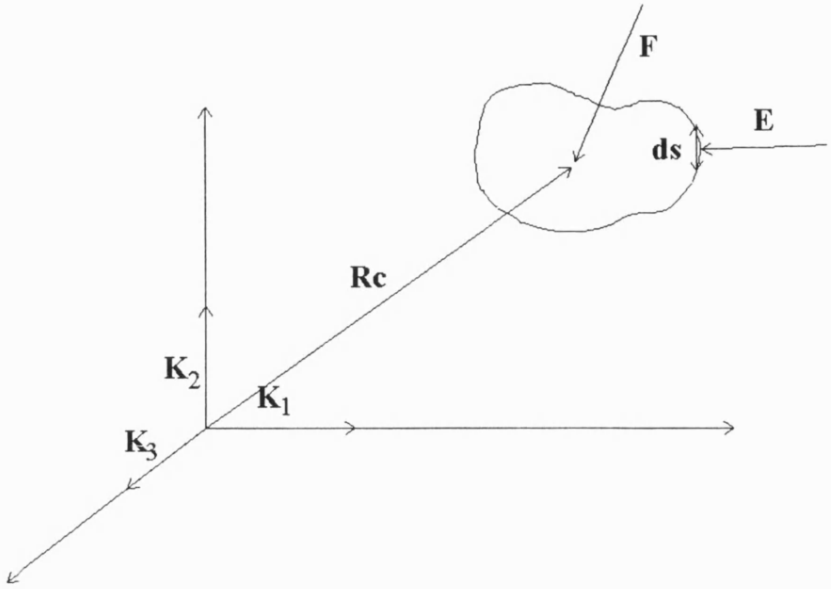


Figure 2.5: Balance of forces for a volume element

There are two forces that act on the element: one is our external (volume) force $\mathbf{F}(x^i,t)$, the other is the total force that the rest of the body exerts on the element through its boundaries. Namely:

$$\mathbf{\Phi} = \frac{1}{dV}\int_{\Omega}\mathbf{F}(x^i,t)\Delta V + \mathbf{E}_{total}(\Omega) \tag{2.23}$$

$\mathbf{E}_{total}(\Omega)$ here is the total force of the rest of the body on the element. It will be shown later that this is an integral over the boundary, with the integrand being a tensor expression, which cannot be easily represented in a vector form. For the time being it is not important. All we need to know is that it is total force transmitted through the boundary, and because the element is strictly internal, it is the force exerted by the rest of the body on the element. We will call it the internal force.

Using the same ideas as for the acceleration we can re-write the last formula as:

$$\mathbf{\Phi} \cong \mathbf{F}(x_c^i,t) + \mathbf{E}_{total}(\Omega) \tag{2.24}$$

We can also note that:

$$dm(\Omega) = \left(\frac{1}{dV}\int_{\Omega}\rho(x^i,t)\Delta V\right)dV \cong \rho(x_c^i,t)dV \tag{2.25}$$

Now we can re-write our basic equation of motion as:

$$\ddot{\mathbf{R}}(x_c^i,t)\rho(x_c^i,t)dV = \mathbf{F}(x_c^i,t)dV + \mathbf{E}_{total}(\Omega) \tag{2.26}$$

or, having divided by the volume, we can write it as:

$$\ddot{\mathbf{R}}(x_c^i,t)\rho(x_c^i,t) = \mathbf{F}(x_c^i,t) + \frac{1}{dV}\mathbf{E}_{total}(\Omega) \tag{2.27}$$

If we now take it to the limit when $dV \to 0$, we will obtain the following equation:

$$\ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \lim_{dV\to 0}\frac{1}{dV}\mathbf{E}_{total}(\Omega) = \mathbf{F}(x^i,t) + \mathbf{E}(x^i,t), \forall x^i \in I, \forall t > 0 \tag{2.28}$$

We obviously have to assume the specified limit exists for every point.

Here $\mathbf{E}(x^i,t)$ is the internal force, which arises purely due to the solid's deformation (as is clear from its derivation) and is an unknown. This equation must be satisfied for every (internal) point at all times and is in fact the law of motion for our solid.

Let us consider the internal force in more detail.

We do not yet know how to find the internal force but we can already list some of its properties, which we can obtain from its derivation:

- It is a volume force, that is, it acts on the volume elements.
- It is purely a reaction force; it arises due to the solid's internal deformation.
- Let us consider two different solids with equal densities, of the same shape, the same mapping function, with the same external force field but with different material properties (a piece of wood and a piece of rubber, for example). We know they will behave differently during our experiment (because of their different elastic properties), therefore, the only term in the equations of motion that may (and, in our case, must) be different for the two solids is the internal force term. From this we can conclude that the internal force must be a function of the elastic material properties of the solid. It provides an elastic response to the deformation, and it depends on the elastic material properties.
- From the above mentioned properties, we can see the following will also be true: if the body is in undeformed state between the times t_1 and t_2 $(t_1 < t_2)$, then the internal force will be identically zero at these times: $\forall t \in]t_1, t_2[$.

- For constant external force fields the internal force has the following property: it is invariant to the body translation and rotates with the body if the body is subjected to rotation. Indeed, the internal force is defined as a limit involving an integral of the force transmitted through the boundary, taken over the surface of a volume element surrounding a point for which the force is calculated. When we translate the body for a constant vector away from its current position, because the external force is constant, we can expect the stress distribution within the body to remain unchanged. Therefore, the forces transmitted through the surface should not change either. Finally, since the boundary shape will remain unchanged, we can conclude the internal force we will get in the limit will also be the same.

  Similarly, if we subject the body to rotation around a point, we can see that the internal forces transmitted through the boundary of a volume element will rotate as well as the boundary itself. This is because they represent the force exerted by the rest of the body on each surface element, and if the body rotates the force will rotate with it. We can, therefore, conclude that the internal elastic force will rotate as well.

It seems, thus far we have achieved significant results relatively quickly. We already have a basic law of motion for our solid; all we need to do now is to add equations for the internal force, which will allow us to find it. Once we know how to find the internal force, we will be able to solve the modelling problem, at least for problems with fixed boundary, without any further considerations.

In fact, in order to give the necessary definitions and introduce the equations for the internal force, we will need to introduce such non-trivial notions as: curvilinear coordinate systems, transformations between the curvilinear coordinate systems, tensors, covariant derivatives and so on.

The remainder of this chapter will be dedicated to finding this internal force. The basic form of the equation of motion (2.28) will not change.

2.5.2 Need for the material curvilinear coordinate system on a solid

Before introducing curvilinear coordinate systems, it seems appropriate to explain why it is necessary to do so in more detail.

We already have a global Cartesian coordinate system K, in which we have expressed the law of motion. All the terms (vectors) in the equation are vectors in K. So the question then is, why is it not possible (or appropriate) to use K?

From one of the properties of the internal elastic force we have described in the previous section, we can see that the elastic force will rotate with the solid when the solid is rotated, and will not change if the solid is translated for a constant vector. One can, therefore, say it will be best to define the elastic force relative to the body itself, and not the global coordinate system. In this case, as the body rotates the force will rotate with it. This can be achieved by introducing a second coordinate system (which we will call the material coordinate system) somehow attached to the solid (moving and translating with the solid as it deforms).

The real question now is, why does this material coordinate system need to be curvilinear? Can it not be Cartesian?

Figure 2.6 below illustrates why we cannot take it to be Cartesian.

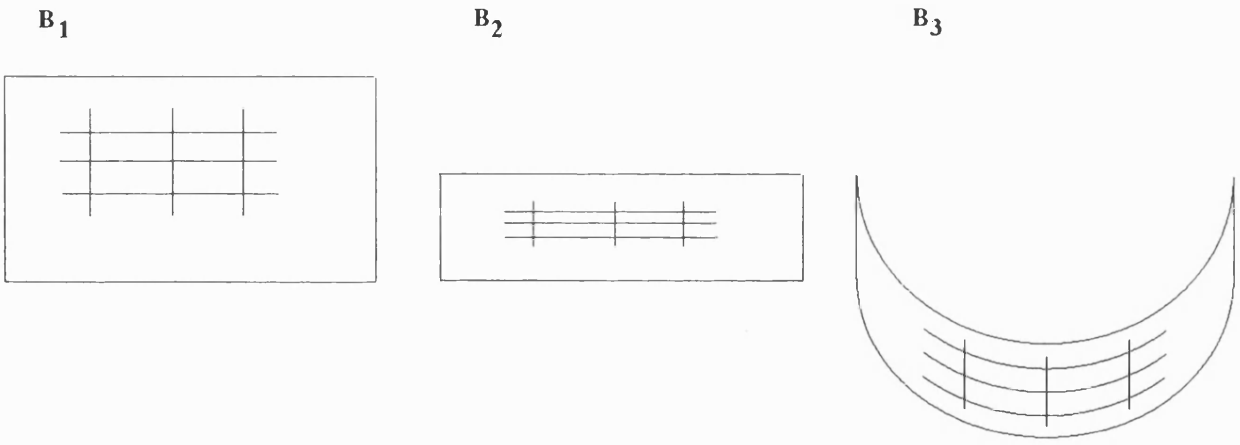


Figure 2.6: Why we need a curvilinear material coordinate system

In this figure we have an undeformed body (B_1), which is first squashed (B_2) and then deformed (B_3) by the external forces.

We can see from this example, why material coordinate system needs to be curvilinear. If it is Cartesian, it will clearly not be very suitable when the body is deformed (B_3).

In fact, our material coordinate system should be curvilinear, and it should deform together with the body itself. So that, in all three cases B_1, B_2 and B_3, the grid shown will correspond to our one (deforming) material coordinate system.

In fact, the need for the curvilinear coordinate system in the classical theory of elasticity is more subtle. As we will show later, in the classical theory of elasticity a notion of a reference frame is introduced, which is normally the solid itself at the start of the experiment. And it is this coordinate system associated with this reference frame, rather than the coordinate system attached to the deforming solid, in which all the quantities at a point are expressed. Therefore, we need to be able to handle the curvilinear coordinate systems in order to be able to deal with curvilinear coordinate systems attached to the reference frames.

Now that we have identified the need for the material coordinate system to be curvilinear, we must introduce the necessary apparatus to be able to work with curvilinear coordinate systems. We must discuss how to express vector quantities in the curvilinear coordinate system, how to differentiate and integrate. Then we will be able to derive formally the expression for the internal elastic force, and write our motion law in the curvilinear material coordinate system.

2.6 Curvilinear Coordinate Systems

2.6.1 Introduction

A note on illustrations: in this chapter, two-dimensional graphical illustrations will be used, as opposed to the full three-dimensional ones. The reason for this is simply a desire to present the concepts in the most descriptive way possible. Since three-dimensional pictures are normally more difficult to understand they will only be used when necessary.

Let us suppose we have a mapping $\mathbf{R}(x^i): I \rightarrow \mathbf{R}^3, \mathbf{R}(x^i) \in C_I^1$, which is differentiable with a continuous derivative. It maps a domain parallelepiped I into a region of space (see Figure 2.7 below).

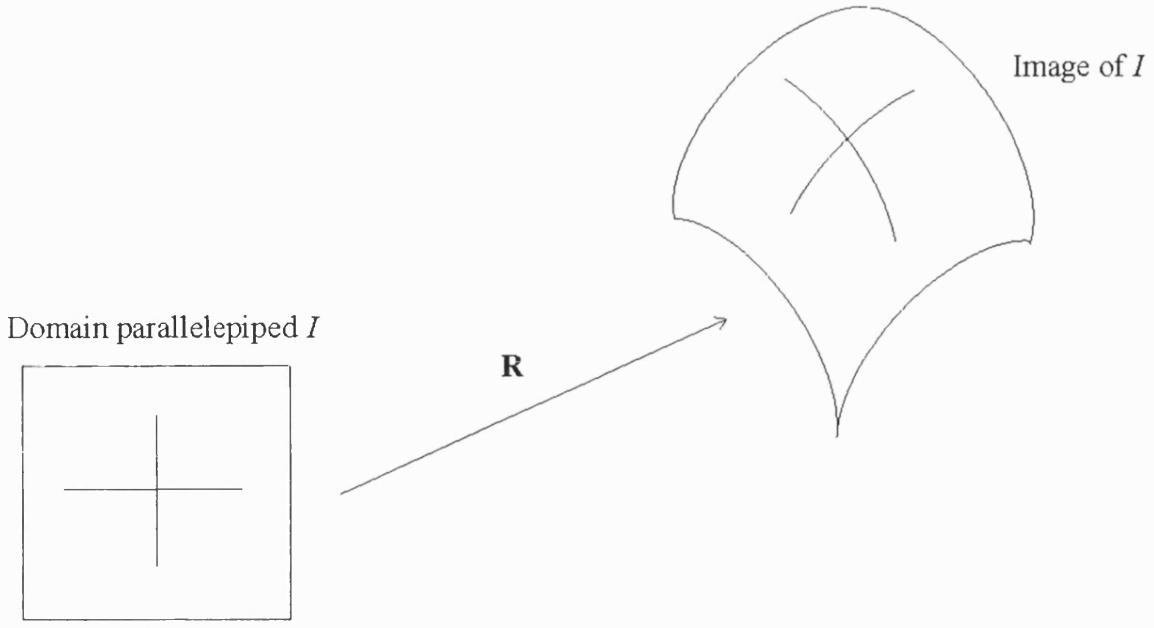


Figure 2.7: Mapping of a domain parallelepiped into $\mathbf{R}^3$

Let us now consider the coordinate lines on the parallelepiped I.

Definition (2.3)

Let $x^i \in I$. Then the set:

$$K^j(x^i) = \left\{ y^i \in I : y^n = x^n, \forall n \neq j \right\} \tag{2.29}$$

is called the j-th coordinate line passing through the point $x^i \in I$. See Figure 2.8 below for an illustration.

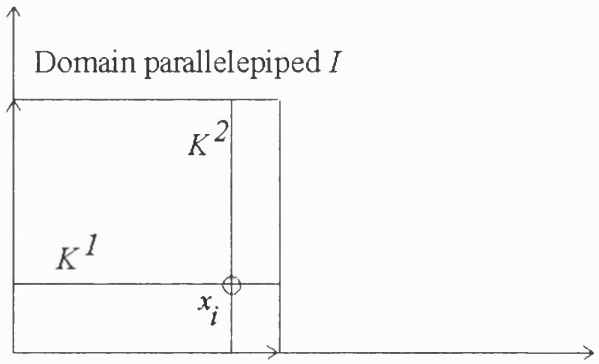


Figure 2.8: Coordinate lines on a domain parallelepiped

Let us now choose a point $x^i \in I$ and fix it, then project the point itself and the coordinate lines passing through it using our mapping function $\mathbf{R}(x^i)$. See a picture Figure 2.9 below for an illustration:

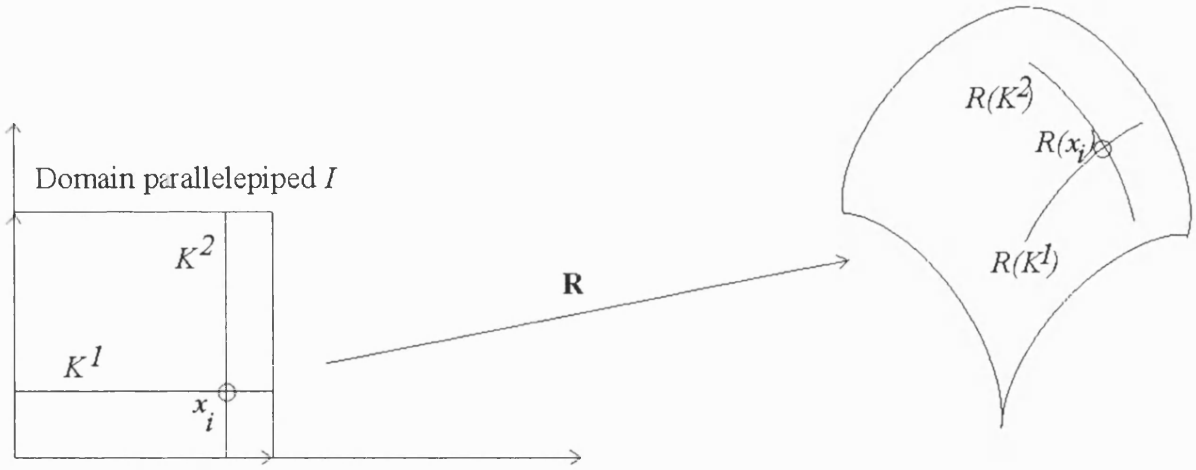


Figure 2.9: Projection of coordinate lines

We now have a coordinate grid in the region of space covered by the mapping $\mathbf{R}(x^i)$ (projections of $K^j(x^i)$). This coordinate grid is said to be induced by the mapping $\mathbf{R}(x^i)$.

In fact, $\mathbf{R}(x^i)$ induces a full curvilinear coordinate system in $\mathbf{R}^3$ (more precisely, in the region of $\mathbf{R}^3$ covered by the mapping $\mathbf{R}(x^i)$).

2.6.2 Covariant base vectors

We are now going to introduce a basic attribute of any coordinate system: base vectors or its basis.

Definition (2.4)

The triple of vectors $\mathbf{G}_i(x^k)$:

$$\mathbf{G}_i(x^k) = \frac{\partial}{\partial x^i}\mathbf{R}(x^k) \qquad (2.30)$$

is called a set of covariant base vectors for the coordinate system induced by $\mathbf{R}(x^i)$, or simply its covariant basis.

We will in the future write $\mathbf{G}_i(x^k)$ in a simpler form, not showing their dependency on x^k (but obviously always assuming it):

$$\mathbf{G}_i = \mathbf{G}_i(x^k) \qquad (2.31)$$

Each of the base vectors $\mathbf{G}_i$ is tangential to the projection of the coordinate line $\mathbf{R}(K^i)$ (see Figure 2.10 below).

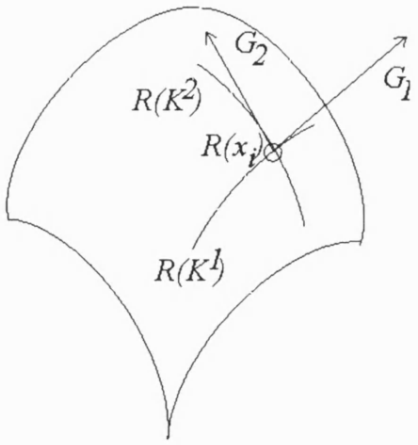


Figure 2.10: Covariant base vectors

This set of base vectors defines a local coordinate system for each point. We can now express any vector quantity defined in the neighbourhood of a point in this new coordinate system, for example a force acting on the point:

$$\mathbf{F} = F^1\mathbf{G}_1 + F^2\mathbf{G}_2 + F^3\mathbf{G}_3 \tag{2.32}$$

Note on notation:

1. From now on we will use the so-called summation convention: whenever two indices appear in the same expression both as upper index and lower index, we assume a sum over the index. For example:

$$\mathbf{F} = F^1\mathbf{G}_1 + F^2\mathbf{G}_2 + F^3\mathbf{G}_3 = \sum_i F^i\mathbf{G}_i = F^i\mathbf{G}_i \tag{2.33}$$

   This is also true for any other expression, for example,

$$a = T^{ik}_j S_{ik} P^j = \sum_i \sum_j \sum_k T^{ik}_j S_{ik} P^j \tag{2.34}$$

2. If in an expression an index is free, that is, it has no corresponding lower or upper index to form a sum, it is assumed the expression is valid for all values of that index. For example,

$$r_j = T_j^{ik} S_{ik} \Leftrightarrow$$
$$r_j = \sum_i \sum_k T_j^{ik} S_{ik}, \forall j = 1,2,3 \tag{2.35}$$

2.6.3 Metric tensor and contravariant base vectors

Let us choose an internal point $x^i \in I$, $\mathbf{R}(x^i)$ is its projection. We also choose a small displacement dx^i from x^i. We will now look at the image **ds** of the displacement dx^i. See Figure 2.11 below:

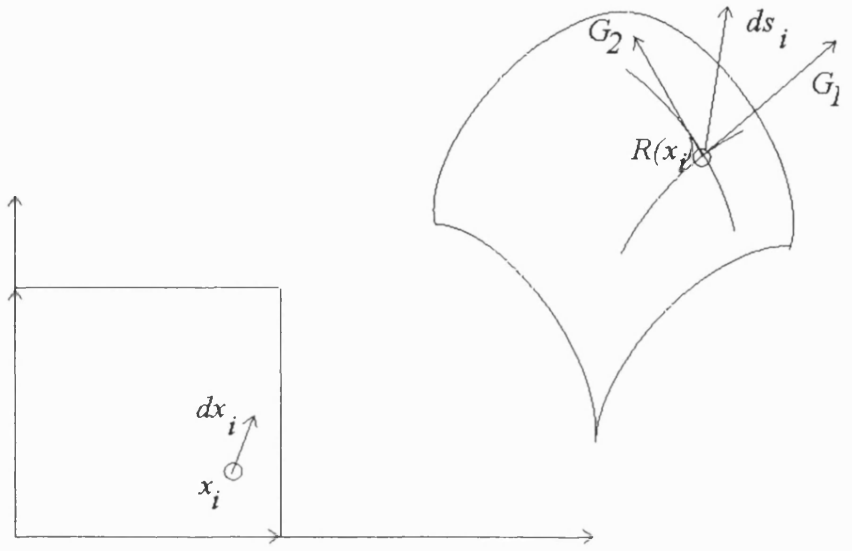


Figure 2.11: Small displacement dx^i and its image ds

If dx^i is small, we can write approximately:

$$\mathbf{ds} = \frac{\partial \mathbf{R}(x^k)}{dx^i} dx^i = \mathbf{G}_i \, dx^i \tag{2.36}$$

We will now consider the square ds^2 of the length of the displacement **ds**:

$$ds^2 = \mathbf{ds} \cdot \mathbf{ds} = \mathbf{G}_i dx^i \cdot \mathbf{G}_j \, dx^j = \mathbf{G}_i \cdot \mathbf{G}_j \, dx^i dx^j \tag{2.37}$$

If we now introduce a new quantity:

$$G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j \tag{2.38}$$

The expression for the square of an element of length ds^i becomes:

$$ds^2 = G_{ij} dx^i dx^j \tag{2.39}$$

Definition (2.5)

G_{ij} is called the metric tensor.

It is metric, because, as we have just shown, it takes part in the expression for the length of the distance between two close neighbouring points. The reason why it is called a tensor will become clear later on, when we formally introduce tensors.

From this definition we can see that the metric tensor is symmetric:

$$G_{ij} = G_{ji} \tag{2.40}$$

Obviously, G_{ij} depends both on x^i, and on the mapping $\mathbf{R}(x^i)$. Let us look at a few specific examples of $\mathbf{R}(x^i)$ to familiarise ourselves with the notions we have introduced so far, and gain a better understanding of the covariant base vectors and metric tensor.

Example 1a: Cartesian coordinate system

First we will look at a trivial example, when $\mathbf{R}(x^i)$ is given simply as:

$$\mathbf{R}(x^i) \equiv x^i \mathbf{k}_i \tag{2.41}$$

That is, a domain parallelepiped I is mapped on itself identically (for any values of A^i). In this case:

$$\mathbf{G}_i(x^k) = \mathbf{k}_i \tag{2.42}$$

where $\mathbf{k}_i$ are unit base vectors of our global coordinate system K.

$$G_{ij} = \mathbf{k}_i \cdot \mathbf{k}_j = \delta_{ij} \tag{2.43}$$

We can see in this example that $\mathbf{R}(x^i)$ induces a Cartesian coordinate system for which the covariant basis coincides with the basis of our global Cartesian coordinate system at any point.

Example 1b:

Let us now look at a slightly modified example, when $\mathbf{R}(x^i)$ is given as:

$$\mathbf{R}(x^i) \equiv (\Xi_i^j x^i + C^j)\mathbf{k}_j$$
$$\Xi_i^j : \sum_s \Xi_i^s \Xi_j^s = \delta_{ij} \tag{2.44}$$

Ξ_i^j here is a rotational matrix (the additional condition specifies that its transposed matrix is its inverse, which in matrix form is $\Xi \cdot \Xi^T = \mathbf{I}$), and C^i is a constant vector.

In this case, our parallelepiped I is first rotated by applying the matrix Ξ_i^j and then displaced by C^i.

We now intend to show that the coordinate system induced by this mapping $\mathbf{R}(x^i)$ is also a Cartesian coordinate system, rotated relatively to our global coordinate system K.

Indeed,

$$\mathbf{G}_s(x^i) = \frac{\partial\left((\Xi_i^j x^i + C^i)\mathbf{k}_j\right)}{dx^s} = \Xi_i^j \delta_s^i \mathbf{k}_j = \Xi_s^j \mathbf{k}_j \tag{2.45}$$

To show $\mathbf{G}_s$ represent a Cartesian coordinate system, we need to show that:

$$G_{ij} = \delta_{ij} \tag{2.46}$$

We have:

$$G_{ij} = \Xi_i^s \mathbf{k}_s \cdot \Xi_j^n \mathbf{k}_n = \Xi_i^s \Xi_j^n \mathbf{k}_s \cdot \mathbf{k}_n = \Xi_i^s \Xi_j^n \delta_{sn} = \sum_s \Xi_i^s \Xi_j^s = \delta_{ij} \tag{2.47}$$

We can see from this example that the metric tensor represents the internal metric properties of the coordinate system induced by the mapping $\mathbf{R}(x^i)$, it is invariant to rotation and translation through a constant vector (as indeed one can expect).

Example 2: Polar coordinate system (two dimensional case)

The polar coordinate system is introduced in the two-dimensional case only (in three dimensions its equivalent is called the spherical coordinate system).

The coordinate system induced by the mapping $\mathbf{R}(x^i) : I^2[A^1, 2\pi] \to \mathbf{R}^2, i = 1,2$:

$$\mathbf{R}(x^i) = \left(C^1 + x^1 \cos(x^2)\right)\mathbf{k}_1 + \left(C^2 + x^1 \sin(x^2)\right)\mathbf{k}_2 \tag{2.48}$$

is called the polar coordinate system, C^1 and C^2 are some constants (see Figure 2.12 below).

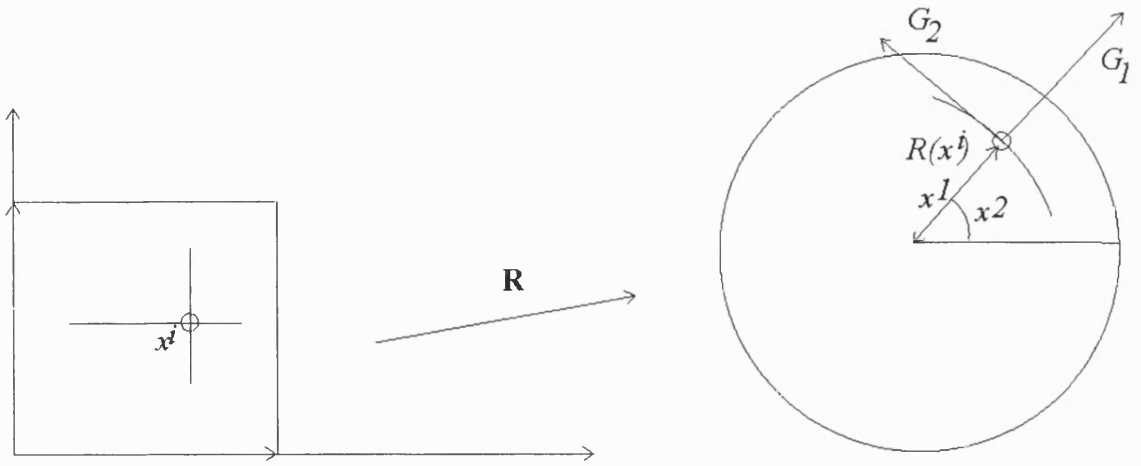


Figure 2.12: Polar coordinate system

The polar coordinate system has a centre with the coordinates $\mathbf{C} = C^1\mathbf{k}_1 + C^2\mathbf{k}_2$. The first coordinate x^1 is the distance of the current point $\mathbf{R}(x^i)$ to this centre $\mathbf{C}$, and the second coordinate x^2 represents the angle between a fixed line passing through the centre $\mathbf{C}$ and the line passing through the current point and the centre $\mathbf{C}$.

Let us find the base vectors and metric tensor for the polar coordinate system:

$$\mathbf{G}_1 = \frac{\partial}{dx^1}\mathbf{R}(x^i) = \cos(x^2)\mathbf{k}_1 + \sin(x^2)\mathbf{k}_2$$

$$\mathbf{G}_2 = \frac{\partial}{dx^2}\mathbf{R}(x^i) = -x^1\sin(x^2)\mathbf{k}_1 + x^1\cos(x^2)\mathbf{k}_2$$

$$G_{11} = 1 \tag{2.49}$$

$$G_{12} = G_{21} = 0$$

$$G_{22} = \left(x^1\right)^2$$

We can see that this coordinate system has some similar metric properties to the Cartesian one:

- Its base vectors $\mathbf{G}_1$ and $\mathbf{G}_2$ are perpendicular to each other (that is, the polar coordinate system has a orthogonal basis);

- The length of the first base vector $\mathbf{G}_1$ is 1.

Normally, in the elasticity literature, the following notation is used for the independent variables of the polar coordinate system:

$$x^1 = r$$
$$x^2 = \varphi \tag{2.50}$$

With this notation, the expressions for the base vectors and the metric tensor will take the following more familiar form:

$$\mathbf{G}_1 = \cos(\varphi)\mathbf{k}_1 + \sin(\varphi)\mathbf{k}_2$$
$$\mathbf{G}_2 = -r\sin(\varphi)\mathbf{k}_1 + r\cos(\varphi)\mathbf{k}_2$$
$$G_{11} = 1 \tag{2.51}$$
$$G_{12} = G_{21} = 0$$
$$G_{22} = r^2$$

It should be noted here that from these formulae it is obvious that the polar coordinate system has a problem at its centre (when $r = 0$). This point is singular and should normally be avoided in practice (that is, it should lie outside the modelled body).

It is well known that in the case of the Cartesian coordinate system, having a triple $\mathbf{k}_i$ of base vectors is sufficient: all vectors can be decomposed in this triple, scalar and vector products can be expressed purely in terms of the coordinates:

$$\mathbf{a} = a^i\mathbf{k}_i$$
$$\mathbf{b} = b^i\mathbf{k}_i$$
$$\mathbf{a}\cdot\mathbf{b} = \Sigma a^i b^i \tag{2.52}$$
$$\mathbf{c} = \mathbf{a}\times\mathbf{b}$$
$$c_1 = a^2b^3 - a^3b^2$$

$$c_2 = -(a^1 b^3 - a^3 b^1)$$

$$c_3 = a^1 b^2 - a^2 b^1$$

In the case of our curvilinear coordinate system induced by $\mathbf{R}(x^i)$ the situation is more complex. Whilst it is still true that any vector can be resolved in base vectors $\mathbf{G}_i$, the expressions (2.52) for the scalar and vector product are generally no longer valid.

Indeed,

$$\mathbf{a} = a^i \mathbf{G}_i$$

$$\mathbf{b} = b^j \mathbf{G}_j \qquad (2.53)$$

$$\mathbf{a} \cdot \mathbf{b} = a^i \mathbf{G}_i b^j \mathbf{G}_j = a^i b^j \mathbf{G}_i \cdot \mathbf{G}_j = a^i b^j G_{ij}$$

Generally, $G_{ij} \neq \delta_{ij}$ therefore this expression for the scalar product is different from the one in the case of the Cartesian coordinate system above.

For a vector product, it can be shown that in the general curvilinear case the expression for the vector product will differ from the one for the Cartesian coordinate system by a scalar multiplication factor. Vector products in curvilinear coordinate systems will be considered in detail later in section 2.6.7.

In fact, in the case of curvilinear coordinate systems we need to introduce an additional (complementary) set of base vectors called contravariant base vectors, or the contravariant basis.

Definition (2.6)

A set of vectors $\mathbf{G}^i = \mathbf{G}^i(x^k)$ such that the following conditions are satisfied:

$$\mathbf{G}^i \cdot \mathbf{G}_j = \delta^i_j \qquad (2.54)$$

is called a set of contravariant base vectors.

From the definition we can see that each $\mathbf{G}^i$ is perpendicular to each $\mathbf{G}_j$ (when $j \neq i$) and its length is such that the scalar product $\mathbf{G}^i \cdot \mathbf{G}_i = 1$ (no summation here).

We can now resolve each vector **a** into components in two different ways:

$$\mathbf{a} = a^i \mathbf{G}_i = a_i \mathbf{G}^i \tag{2.55}$$

The components a^i, that is, the components of vector **a** in the covariant basis, are called the contravariant components of the vector.

Similarly, the components a_i, that is, the components of vector **a** in the contravariant basis, are called the covariant components of the vector.

To memorise this naming convention, it is useful to remember the following rule: any quantity (vector or scalar) that has an upper index is contravariant, and if it has a lower index, it is covariant.

In the light of this new definition, let us now see how we can utilise the contravariant base vectors (and covariant vector components) in the expression for a scalar product:

$$\mathbf{a} = a^i \mathbf{G}_i$$

$$\mathbf{b} = b_j \mathbf{G}^j \tag{2.56}$$

$$\mathbf{a} \cdot \mathbf{b} = a^i \mathbf{G}_i b_j \mathbf{G}^j = a^i b_j \mathbf{G}_i \cdot \mathbf{G}^j = a^i b_j \delta^i_j = a^i b_i$$

Similarly,

$$\mathbf{a} \cdot \mathbf{b} = a_i b^i \tag{2.57}$$

Now we can see that this expression for the scalar product is similar to the one in the Cartesian coordinate system. In fact, it is not difficult to see that in the case of the Cartesian coordinate system, the covariant and contravariant base vectors are the same:

$$\mathbf{k}^i = \mathbf{k}_i, \forall i = 1,2,3 \text{ (no summation is assumed)} \tag{2.58}$$

Therefore, the covariant and contravariant coordinates for any vector coincide and the expression shown above for the scalar product in the curvilinear coordinate system is a general expression, including the one we have shown earlier for the Cartesian coordinate system as a special case.

We have introduced a metric tensor:

$$G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j \tag{2.59}$$

In fact, we can now say that it is a definition of the metric tensor components in the covariant form. There is another representation of the metric tensor in the contravariant form.

Definition (2.7)

The metric tensor in its contravariant form is given by the following expression:

$$G^{ij} = \mathbf{G}^i \cdot \mathbf{G}^j \tag{2.60}$$

The name "contravariant" here is used in similarity to the vector case, when one vector can be resolved into covariant and contravariant components.

We can see from this definition that the metric tensor is symmetric in both its covariant and contravariant forms.

From the definition of the contravariant basis, we have:

$$\mathbf{G}^i \cdot \mathbf{G}_j = \delta^i_j \tag{2.61}$$

This expression is valid in any coordinate system. We will see later, after we have formally introduced tensors, that G^{ij}, δ^i_j and G_{ij} are in fact the different formal representations of one single quantity called the metric tensor (namely, contravariant, mixed variance and covariant forms).

Let us now see how the two basis sets $\mathbf{G}_i$ and $\mathbf{G}^i$ are connected to each other. We can show that the metric tensor plays a part in the relationship. Indeed, we can formally resolve each base vector of the contravariant basis in the covariant basis:

$$\mathbf{G}^i = A^{ij}\mathbf{G}_j \tag{2.62}$$

The set of values A^{ij} provides a connection between the two sets of base vectors. We can write,

$$\mathbf{G}^i \cdot \mathbf{G}^k = G^{ik} = A^{ij}\mathbf{G}_j \cdot \mathbf{G}^k = A^{ij}\delta^k_j = A^{ik} \tag{2.63}$$

or,

$$\mathbf{G}^i = G^{ij}\mathbf{G}_j \tag{2.64}$$

similarly,

$$\mathbf{G}_i = G_{ij}\mathbf{G}^j \tag{2.65}$$

We can see that the metric tensor can be used to lower or raise the index of the basis vectors (in the sense given by these expressions).

Another important property of the metric tensor can be obtained if we write:

$$\mathbf{G}_i \cdot \mathbf{G}^k = \delta_i^k = G_{ij}\mathbf{G}^j \cdot \mathbf{G}^k = G_{ij}G^{jk} \tag{2.66}$$

or,

$$G_{ij}G^{jk} = \delta_i^k \tag{2.67}$$

In a matrix form, it means the two matrixes, formed of the covariant and contravariant components of the metric tensor are reciprocal.

Having seen that the metric tensor lowers or raises the index in the base vectors, the natural question will be to investigate if a similar expression can be derived for components of a vector:

$$\begin{aligned} \mathbf{a} &= a^i\mathbf{G}_i = a_i\mathbf{G}^i \\ a_i\mathbf{G}^i &= a^i\mathbf{G}_i = a^iG_{ik}\mathbf{G}^k \end{aligned} \tag{2.68}$$

Or renaming the indices:

$$\left(a_i - a^kG_{ki}\right)\mathbf{G}^i = 0 \tag{2.69}$$

We cannot just cancel the term $\mathbf{G}^i$. But we can take a scalar product with a $\mathbf{G}_j$:

$$\left(a_i - a^kG_{ki}\right)\mathbf{G}^i \cdot \mathbf{G}_j = \left(a_i - a^kG_{ki}\right)\delta_j^i = \left(a_j - a^kG_{kj}\right) = 0 \tag{2.70}$$

We can see from this expression that indeed the metric tensor also lowers and raises the indices of vector components, in the same way as it does in the case of the base vectors:

$$a_i = a^k G_{ki}$$
$$a^i = a_k G^{ki} \tag{2.71}$$

2.6.4 Coordinate transformations

We have so far introduced the full set of tools for curvilinear coordinate systems, given by the mapping **R**. The question we will answer in this section is, what will happen when we have a different mapping $\mathbf{R}'$? How are all the quantities we have for each of the coordinate systems connected?

There are two good reasons to consider coordinate transformations:

- As we have shown earlier, for a body B we can have many mappings, and therefore many different coordinate systems on the same body B.
- More importantly, whatever the coordinate system we have chosen for our body in its original (undeformed) state, it will deform with the body as the body undergoes deformation during our experiment.

Let us have two mappings $\mathbf{R}$ and $\mathbf{R}'$ for the body B covering the same region in space:

$$\mathbf{R}(x^i): I[A^1, A^2, A^3] \rightarrow \mathrm{R}^3$$
$$\mathbf{R}'(x^{i'}): I'[A^{1'}, A^{2'}, A^{3'}] \rightarrow \mathrm{R}^3 \tag{2.72}$$

Figure 2.13 below gives a graphical illustration of this.

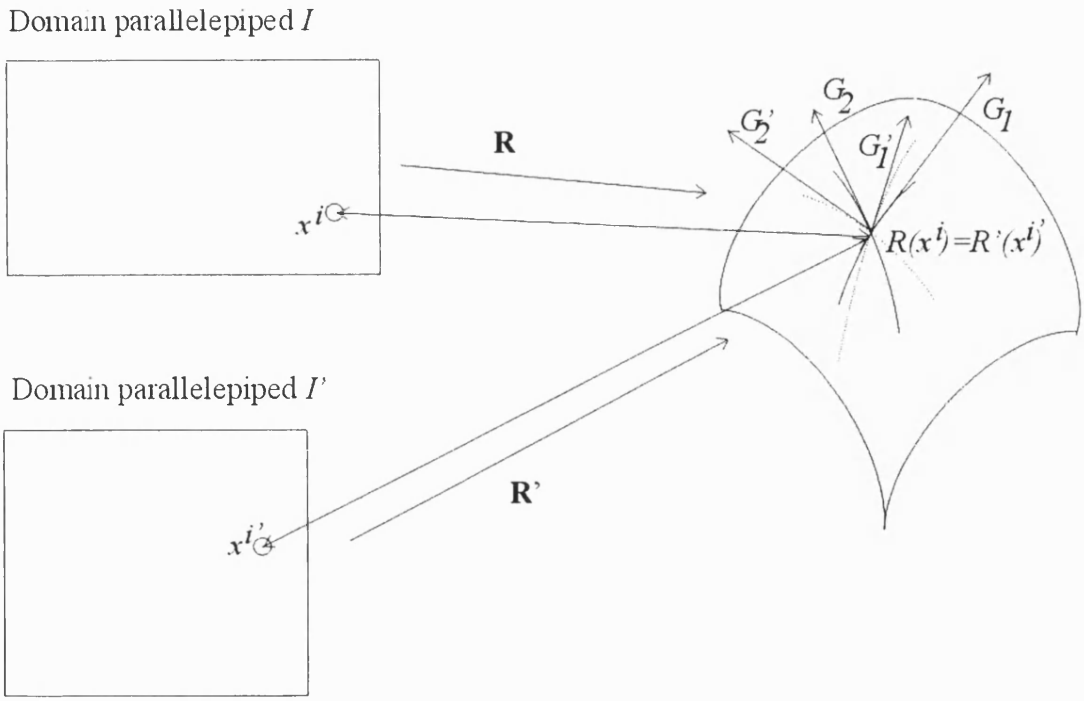


Figure 2.13: Two different mappings for the same body B

We can see $x^{i'}$ and x^i are bound by a one-to-one relationship (mapping):

$$x^{i'} = x^{i'}(x^k)$$
$$x^{i'} = \mathbf{R}'^{-1}(\mathbf{R}(x^k)) \tag{2.73}$$

Now we consider a small displacement **ds** (away from a point $\mathbf{R}(x^k)$):

$$\mathbf{ds} = \mathbf{G}_i\, dx^i = \mathbf{G}'_{i'}\, dx^{i'} \tag{2.74}$$

But from the previous expression giving dependency of $x^{i'}$ and x^i, we can write:

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^k} dx^k \tag{2.75}$$

We will denote the coefficient of dx^k as:

$$\beta_k^{i'} = \frac{\partial x^{i'}}{\partial x^k} \tag{2.76}$$

We can then write:

$$dx^{i'} = \beta_k^{i'} dx^k \tag{2.77}$$

Now using the expression for **ds** we can write the following equality:

$$\mathbf{G}_i dx^i = \mathbf{G}'_{i'} \beta_k^{i'} dx^k \tag{2.78}$$

If we now change the summation index on the right, move the term on the right to the left, and take the common coefficient out, we will obtain the following equality:

$$\left(\mathbf{G}_i - \mathbf{G}'_{i'} \beta_i^{i'}\right) dx^i = 0 \tag{2.79}$$

Again, we cannot simply cancel the term dx^i here (because it is a sum). We can use the following technique: **ds** is arbitrary, so we choose it so that $dx^1 > 0$, $dx^2 = 0$ and $dx^3 = 0$. From this special case, we can conclude that:

$$\begin{aligned} &\left(\mathbf{G}_1 - \mathbf{G}'_{i'} \beta_1^{i'}\right) dx^1 = 0 \Rightarrow \\ &\mathbf{G}_1 = \mathbf{G}'_{i'} \beta_1^{i'} \end{aligned} \tag{2.80}$$

If we now continue the process and take other individual displacements to be zero in turn, we can conclude that in general:

$$\mathbf{G}_i = \mathbf{G}'_{i'} \beta_i^{i'} \tag{2.81}$$

Similarly, we can also derive the following equality for the contravariant base vectors:

$$\mathbf{G}^i = \mathbf{G}'^{i'} \beta_{i'}^i \tag{2.82}$$

Where,

$$\beta_i^{i'} = \frac{\partial x^{i'}}{\partial x^i} \tag{2.83}$$

Using these transformation rules for the base vectors we can obtain the transformation formulae for the components of a vector,

$$\begin{aligned} &a_i = \beta_i^{i'} a_{i'} \Leftrightarrow a_{i'} = \beta_{i'}^i a_i \\ &a^i = \beta_{i'}^i a^{i'} \Leftrightarrow a^{i'} = \beta_i^{i'} a^i \end{aligned} \tag{2.84}$$

as well as for the metric tensor,

$$G_{ij} = \beta_i^{i'}\beta_j^{j'}G_{i'j'} \Leftrightarrow G_{i'j'} = \beta_{i'}^{i}\beta_{j'}^{j}G_{ij}$$
$$G^{ij} = \beta_{i'}^{i}\beta_{j'}^{j}G^{i'j'} \Leftrightarrow G^{i'j'} = \beta_i^{i'}\beta_j^{j'}G^{ij} \tag{2.85}$$

We can note from these expressions that it takes one β for each index when a quantity's coordinates are transformed from one coordinate system into another one.

We conclude this section with an example.

Coordinate transformation between a two-dimensional Cartesian coordinate system and a two-dimensional polar coordinate system

We will present the quantities in the Cartesian coordinate system as unstressed, and the ones in the polar coordinate system as stressed.

We have already calculated:

$$\mathbf{G}_i = \mathbf{k}_i$$
$$G_{ij} = \delta_{ij}$$
$$G^{ij} = \delta^{ij}$$
$$\mathbf{G}'_{1'} = \sin(x^{2'})\mathbf{k}_{1'} + \cos(x^{2'})\mathbf{k}_{2'}$$
$$\mathbf{G}'_{1'} = x^{1'}\cos(x^{2'})\mathbf{k}_{1'} - x^{1'}\sin(x^{2'})\mathbf{k}_{2'} \tag{2.86}$$
$$G'_{1'1'} = G'^{1'1'} = 1$$
$$G'_{2'1'} = G'_{1'2'} = G'^{2'1'} = G'^{1'2'} = 0$$
$$G'_{2'2'} = \left(x^{1'}\right)^2$$

In order to obtain transformation formulae, we will need a direct mapping,

$$x^{i'} = x^{i'}(x^k) \tag{2.87}$$

and the reverse,

$$x^k = x^k(x^{i'}) \tag{2.88}$$

These mappings can be easily found:

$$x^1 = x^{1'}\cos(x^{2'}),\ x^2 = x^{1'}\sin(x^{2'}) \tag{2.89}$$

and the reverse:

$$x^{1'} = \sqrt{(x^1)^2 + (x^2)^2},\ x^{2'} = \cot\frac{x^1}{x^2} \tag{2.90}$$

Or if we use a more familiar notation: $x^1 = x$, $x^2 = y$ and $x^{1'} = r$, $x^{2'} = \varphi$ these mapping relationships will take the following form:

$$\begin{gathered} x = r\cos(\varphi),\ y = r\sin(\varphi) \\ r = \sqrt{x^2 + y^2},\ \varphi = \cot\frac{x}{y} \end{gathered} \tag{2.91}$$

Now we can easily find the transformation coefficients:

$$\begin{gathered} \beta_x^r = \frac{x}{\sqrt{x^2 + y^2}} = \frac{r\cos(\varphi)}{r} = \cos(\varphi) \\ \beta_y^r = \frac{y}{\sqrt{x^2 + y^2}} = \frac{r\sin(\varphi)}{r} = \sin(\varphi) \\ \beta_x^\varphi = \frac{y}{x^2 + y^2} = \frac{r\sin(\varphi)}{r^2} = \frac{\sin(\varphi)}{r} \\ \beta_y^\varphi = -\frac{x}{x^2 + y^2} = -\frac{r\cos(\varphi)}{r^2} = -\frac{\cos(\varphi)}{r} \end{gathered} \tag{2.92}$$

Now let us suppose we have a vector **a** given in the Cartesian coordinate system:

$$\mathbf{a} = a^i\mathbf{G}_i = a^i\mathbf{k}_i \tag{2.93}$$

We want to find out how these coordinates a^i will transform into the polar coordinate system. For this we use the transformation rule we have just devised:

$$a^r = \beta_x^r a^x + \beta_y^r a^y \tag{2.94}$$

$$a^{\varphi} = \beta_x^{\varphi} a^x + \beta_y^{\varphi} a^y$$

Or, utilising the formulae we have just obtained:

$$\begin{aligned} a^r &= \cos(\varphi) a^x + \sin(\varphi) a^y = \frac{x}{\sqrt{x^2+y^2}} a^x + \frac{y}{\sqrt{x^2+y^2}} a^y \\ a^r &= \frac{\sin(\varphi)}{r} a^x - \frac{\cos(\varphi)}{r} a^y = \frac{y}{x^2+y^2} a^x - \frac{x}{x^2+y^2} a^y \end{aligned} \tag{2.95}$$

Similarly, we can find the reverse transformation from the polar to Cartesian coordinate system.

Note that the transformation expressions can be given in either coordinate system's coordinates: x, y or r, φ (because β can be expressed in either coordinates).

2.6.5 Tensors

The tensor concept plays an important part in continuum mechanics and physics. It has emerged from the need to express physical quantities distributed over a body or a space region as well as expressions involving those quantities, in a way that will preserve their form whatever the underlying coordinate system.

Let us consider a simple example. If we have a vector in a curvilinear coordinate system with a basis $\mathbf{G}_i$:

$$\mathbf{a} = a^i \mathbf{G}_i \tag{2.96}$$

This vector may represent a certain physical quantity, it is not important for this example.

We know how this vector's coordinates will transform into another coordinate system $\mathbf{G}'_{i'}$:

$$\begin{aligned} \mathbf{a} &= a'^{i'} \mathbf{G}'_{i'} \\ a'^{i'} &= \beta_i^{i'} a^i \end{aligned} \tag{2.97}$$

Let us now suppose we have somehow derived that:

$$a^i = 0 \tag{2.98}$$

That is, we have somehow concluded that in one coordinate system $\mathbf{G}_i$ vector a^i is such that its components are all zero.

Now, by using the transformation law, we can see that in any other coordinate system we also have:

$$a'^{i'} = \beta_i^{i'} a^i = 0 \qquad (2.99)$$

That is, having derived equality $a^i = 0$ for one coordinate system, because a^i is a vector, we can immediately state that the equation is valid for any other coordinate system.

From this derivation it follows that any vector equation stated in one coordinate system is true in any other. Indeed, if we have, for example:

$$b^i + De^i + f^i = g^i \qquad (2.100)$$

We can transform it into a form:

$$a^i = 0$$

where,

$$a^i = b^i + De^i + f^i - g^i$$

In fact, physically, this conclusion is fairly obvious. Indeed, we should remember that a^i is supposed to be a mathematical abstraction corresponding to a real physical quantity, for example force. If we have a zero force field for a body or a space region, it is quite clear that the mathematical entity that is to correspond to the force should also be zero and it should not depend on the coordinate system in which the mathematical quantity is expressed.

This property of vector equations to have the same form for all coordinate system is at the heart of the tensor concept. We are now going to introduce tensors formally.

Definition (first order tensor) *(2.8)*

Any vector a^i is called a tensor of the first order (or rank). As before, we say that a^i are the contravariant components of the tensor and a_i are the covariant components of the same tensor. We should note here that the order number refers to the number of indices in the quantity.

Vectors are not sufficient to represent all physical quantities. We have already seen the metric tensor G_{ij} : it has 9 components, and it is clear that it will be impossible to represent it as a three dimensional vector (a pointed arrow). It can be formally considered a vector in the nine-dimensional space, but this consideration is very formal and has little physical meaning.

Definition (contravariant components of a tensor) *(2.9)*

Quantity A^{ij} is called a set of contravariant components of the tensor of a second order, if it transforms between two coordinate systems according to the following rule:

$$A'^{i'j'} = \beta_i^{i'} \beta_j^{j'} A^{ij} \tag{2.101}$$

The contravariant components of the second order tensors are said to show double-contravariant behaviour when they are transformed from one coordinate system into another.

Definition (covariant components of a tensor) *(2.10)*

Quantity A_{ij} is called a set of covariant components of a tensor of a second order, if it transforms between two coordinate systems according to the following rule:

$$A'_{i'j'} = \beta_{i'}^{i} \beta_{j'}^{j} A_{ij} \tag{2.102}$$

Similarly to the contravariant case, the covariant components of the second order tensors are said to show double-covariant behaviour when they are transformed from one coordinate system into another.

Definition (mixed variance components of a tensor) *(2.11)*

The following quantities are called the components of mixed variance of a second order tensor:

$$\begin{aligned} A'^{\cdot j'}_{i'} &= \beta_{i'}^{i} \beta_{j}^{j'} A^{\cdot j}_{i} \\ A'^{i'}_{\cdot j'} &= \beta_{i}^{i'} \beta_{j'}^{j} A^{i}_{\cdot j} \end{aligned} \tag{2.103}$$

It must be noted that generally $A^{\cdot j}_{i}$ and $A^{i}_{\cdot j}$ are different (hence the use of dots to show which index is in which position).

We have referred in all these definitions to various types of components of a tensor: contravariant, covariant and mixed variance. We now need to show that indeed A^{ij}, A_{ij}, $A_{i\cdot}^{\cdot j}$ and $A_{\cdot j}^{i\cdot}$ all refer to the same object, which is a tensor of the second order.

We can state that there exist vectors a_i, b_i, d_i and e_i such that:

$$A_{ij} = a_i b_j + d_i e_j \tag{2.104}$$

We have to consider four vectors, just two $A_{ij} = a_i b_j$ will not be generally sufficient: two vectors will produce a maximum of six independent products $a_i b_j$, whilst A_{ij} can have nine.

Let us now consider a quantity $A_{ij}G^{ik}$:

$$A_{ij}G^{ik} = a_i G^{ik} b_j + d_i G^{ik} e_j = a^k b_j + d^k e_j = A_j^{k\cdot} \tag{2.105}$$

We can see that we can use G^{ik} to lower one index of a tensor. Similarly, the following expressions can be obtained:

$$\begin{aligned} A_{ij}G^{jk} &= A_{i\cdot}^{\cdot k} \\ A_{ij}G^{ik}G^{jl} &= A^{kl} \end{aligned} \tag{2.106}$$

We can see from these equalities, that indeed A^{ij}, A_{ij}, $A_{i\cdot}^{\cdot j}$ and $A_{\cdot j}^{i\cdot}$ are all representations of the same object, a second order tensor, given in different forms. It is similar to vectors (with covariant and contravariant components) but more complex because there are two independent indices.

A tensor for which the following equality holds:

$$\begin{aligned} A^{ij} &= A^{ji} \\ A_{ij} &= A_{ji} \end{aligned} \tag{2.107}$$

is called a *symmetric* tensor. One can easily see that one of these conditions is enough for the other to follow, that is, $A^{ij} = A^{ji} \Rightarrow A_{ij} = A_{ji}$ and $A_{ij} = A_{ji} \Rightarrow A^{ij} = A^{ji}$.

If a tensor is symmetric, we can see that:

$$A_{ij}G^{ik} = A^{k\cdot}_{\cdot j} = A_{ji}G^{ik} = A^{\cdot k}_{j\cdot} \tag{2.108}$$

Or in short:

$$A^{k\cdot}_{\cdot j} = A^{\cdot k}_{j\cdot} \tag{2.109}$$

We can see from this equality that for symmetric tensors there is only one set of components of mixed variance, and we will simply write $A^k_j = A^{k\cdot}_{\cdot j} = A^{\cdot k}_{j\cdot}$.

One second order tensor has already been introduced before: the metric tensor G_{ij}. From its definition we know it is symmetric, and also that in mixed variance forms it is simply the Cronecker delta:

$$G_{ij}G^{ik} = G^k_j = \delta^k_j \tag{2.110}$$

Let us now define tensors of orders other than one and two.

Definition (tensor of zero-th order) (2.12)

Any scalar quantity, constant in any coordinate system, is formally called a tensor of the zero-th order.

A physical example of such a quantity is material density ρ. It is defined for every point belonging to the body and is the same in any coordinate system (because it is a property of the physical material point to which the geometrical point corresponds to).

There also exist tensors of the orders higher than two. A formal definition of a tensor of order n is quite cumbersome, and we will limit ourselves to order three and covariant components. Tensors of a higher order can be introduced by the analogy.

Definition (covariant components of a tensor of third order) (2.13)

Quantity A_{ijk} is called a set of covariant components of a tensor of third order, if it transforms between two coordinate systems according to the following rule:

$$A'_{i'j'k'} = \beta^i_{i'}\beta^j_{j'}\beta^k_{k'}A_{ijk} \tag{2.111}$$

Tensors of higher order can also be symmetric in a pair of their indices, for example the tensor:

$$A_{ijk} = A_{ikj} \tag{2.112}$$

is symmetric in its second and third indices. This obviously does not presume that there is any symmetry in any other pairs: the first and the second or the first and the third.

We can form new tensors by forming so called tensor products. If we have tensors $A_{ij\cdot}^{\cdot\cdot k}$ and B^{ls}:

$$C_{ij\cdots}^{\cdot\cdot kls} = A_{ij\cdot}^{\cdot\cdot k} B^{ls} \tag{2.113}$$

$C_{ij\cdots}^{\cdot\cdot kls}$ is the tensor product of the two tensors and is itself a tensor of the fifth order (3 + 2).

Whilst tensor products produce tensors of higher order, there is an operation which lowers the order. It is called tensor contraction.

Definition (tensor contraction) *(2.14)*

An operation on a tensor by which two of its indices (one lower covariant and one upper contravariant) are made equal to invoke the summation convention is called contraction.

After a tensor is contracted by two of its indices its order is reduced by two. Let us consider for example the above-mentioned tensor $C_{ij\cdots}^{\cdot\cdot kls}$ and do the contraction on two of its indices i and s:

$$C_{ij\cdots}^{\cdot\cdot kli} = D_{j\cdot\cdot}^{\cdot kl} \tag{2.114}$$

The result tensor $D_{j\cdot\cdot}^{\cdot kl}$ is a third order tensor.

Let us consider the metric tensor $G_j^i = \delta_j^i$ in it mixed variance form. If we form its contraction, we can see we have no free indices left and the result should be a tensor of the zero-th order:

$$G_i^i = \delta_i^i = 3 \tag{2.115}$$

That is, the contraction of the metric tensor produces a scalar 3.

Let us now return to the idea introduced at the beginning of this section. We have seen for vector equations that if we can show they are valid in one coordinate system, they are

automatically valid in any other coordinate system. Vectors are tensors of the first order, and the question now is, is the same statement true for general tensor equations?

The answer is yes. To show this we will first note that any tensor equation is a sum of terms of the same variance equal to zero. Each term can consist of tensor products of other tensors and possible contractions. But altogether, the sum will be a tensor that is equal to zero. For example, we have the following equation:

$$A^{i\cdot k}_{\cdot j\cdot} B_k + C^{i\cdot l}_{\cdot jl} = D^{i\cdot}_{\cdot j} \Leftrightarrow$$
$$E^{i\cdot}_{\cdot j} = A^{i\cdot k}_{\cdot j\cdot} B_k + C^{i\cdot l}_{\cdot jl} - D^{i\cdot}_{\cdot j} = 0 \tag{2.116}$$

Therefore, we can say that a general tensor equation will be of the form where a tensor of mixed variance is equal to zero. When the tensor is transformed into another coordinate system, the result will be a sum of product of components of this tensor and the transformation term β. Since each component of the tensor in the old coordinate system is zero, this sum (the new components) will also be zero.

For example, consider an equation:

$$E^{i\cdot}_{\cdot j} = 0 \tag{2.117}$$

which is true in one coordinate system.

When $E^{i\cdot}_{\cdot j}$ is transformed into another coordinate system, we will have the following dependency of the new components over the old ones:

$$E'^{i'\cdot}_{\cdot j'} = \beta^{i'}_{i} \beta^{j}_{j'} E^{i\cdot}_{\cdot j} \tag{2.118}$$

and because each $E^{i\cdot}_{\cdot j} = 0$, we can conclude that $E'^{i'\cdot}_{\cdot j'} = 0$ as well.

Therefore, we can state an important general property of tensors: any tensor equation that is shown to be valid in one coordinate system, will automatically be valid in any other coordinate system.

Again, this property simply reflects the physical reality. A tensor equation states a law that combines various physical quantities. Because this law involves mathematical equivalents of the real physical quantities, and describes a physical law valid for a given set of circumstances (for example for every small part of a body), it should have the same form

regardless of the underlying coordinate system in which the quantities are measured or presumed to be expressed.

In practice, whenever we obtain a quantity that has a number of indices, in order to use the quantity in an equation, we will need to check that this quantity shows tensor character, that is, it represents components of a tensor of given variance.

We have by now introduced most of the major notions on the way to our target, which is to formulate the expression for the elastic force.

We have introduced curvilinear coordinate systems, two sets of base vectors necessary for their effective handling and the metric tensor – the basic quantity holding the important characteristics of the metrical properties of the coordinate system. We have discussed the transformations between the coordinate systems and introduced tensors of arbitrary orders.

There are several issues that still need to be addressed but all of them are based on extending the apparatus introduced so far. We will need to enrich it with the operations and the quantities we will need before we can start introducing the tensor quantities directly involved in the derivation of the elastic force for a material element.

First of all, we need to discuss how to express elementary surface elements in tensor form. This will be needed when we deal with the internal forces inside the body. The next two sections 2.6.6 and 2.6.7 are dedicated to this.

The other remaining issue is differentiation (and integration) – the rate of change. We will need to discuss how to differentiate tensors. This is shown in section 2.6.8. In this section we will also briefly touch upon integration.

2.6.6 Permutation tensor

In some texts the permutation tensor is also called Levi-Civita tensor.

Definition (2.15)

The numbers:

$$\in_{ijk} = \mathbf{G}_i \cdot \left(\mathbf{G}_j \times \mathbf{G}_k \right) \tag{2.119}$$

are the components of the so called permutation tensor.

The fact that it is a tensor is immediately apparent:

$$\mathbf{G}'_{i'} = \beta^{i}_{i'}\mathbf{G}_i \Rightarrow$$

$$\in'_{i'j'k'} = \mathbf{G}'_{i'} \cdot \left(\mathbf{G}'_{j'} \times \mathbf{G}'_{k'}\right) = \beta^{i}_{i'}\beta^{j}_{j'}\beta^{k}_{k'}\mathbf{G}_i \cdot \left(\mathbf{G}_j \times \mathbf{G}_k\right) = \beta^{i}_{i'}\beta^{j}_{j'}\beta^{k}_{k'} \in_{ijk} \tag{2.120}$$

At this stage we have the permutation tensor given only by its covariant components. Before we introduce its contravariant components, it is useful to derive an alternative expression for the permutation tensor.

Definition (permutation symbols) (2.16)

The numbers given by the following definition:

$$e_{ijk} = +1, \text{ if } (i,j,k) \text{ form an even permutation;}$$

$$e_{ijk} = -1, \text{ if } (i,j,k) \text{ form an odd permutation;} \tag{2.121}$$

$$e_{ijk} = 0, \text{ if two or more indices of the sequence } (i,j,k) \text{ are equal;}$$

are called the permutation symbols.

One permutation is a single exchange of two numbers in a given sequence. For example:

$$(1,2,3) \rightarrow (1,3,2)$$

A sequence of ordered triples (i,j,k) is called *even*, if the total number of single permutations needed to transform it to the form $(1,2,3)$ is even, that is, it is one of the following: $(1,2,3)$, $(2,3,1)$ or $(3,1,2)$.

A sequence of ordered triples (i,j,k) is called *odd*, if the total number of single permutations needed to transform it to the form $(1,2,3)$ is odd, that is, it is one of the following: $(3,2,1)$, $(1,3,2)$ or $(2,1,3)$.

If the coordinate system is Cartesian, the permutation tensor will coincide with the permutation symbols:

$$\in_{ijk} = e_{ijk} \tag{2.122}$$

In the general case, from the definition of the permutation tensor and the fact that it is a tensor, we can also derive formally that the permutation tensor must have a similar form to

the permutation symbols and differ in the general form by a factor, which should depend on the coordinate system:

$$\in_{ijk} = Ae_{ijk} \tag{2.123}$$

In the Cartesian coordinate system it is obvious that $A = 1$.

It can be shown (see for example Flugge [91] for an excellent detailed proof) that indeed if we consider:

$$G_{ij} = \begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix} \tag{2.124}$$

and also:

$$G = \det\left(G_{ij}\right) \tag{2.125}$$

Then $A = \sqrt{G}$, that is, the alternative expression for the permutation tensor in the covariant form will be:

$$\begin{aligned} &\in_{ijk} = +\sqrt{G}, \text{ if } (i,j,k) \text{ form an even permutation;} \\ &\in_{ijk} = -\sqrt{G}, \text{ if } (i,j,k) \text{ form an odd permutation;} \\ &\in_{ijk} = 0, \text{ if two or more indices of the sequence } (i,j,k) \text{ are equal;} \end{aligned} \tag{2.126}$$

It can also be shown that the contravariant components of the permutation tensor are given by:

$$\begin{aligned} &\in^{ijk} = +\frac{1}{\sqrt{G}}, \text{ if } (i,j,k) \text{ form an even permutation;} \\ &\in^{ijk} = -\frac{1}{\sqrt{G}}, \text{ if } (i,j,k) \text{ form an odd permutation;} \\ &\in^{ijk} = 0, \text{ if two or more indices of the sequence } (i,j,k) \text{ are equal;} \end{aligned} \tag{2.127}$$

Indeed, we can use the metric tensor G^{ij} to raise all three indices of the permutation tensor given by its covariant components $\in_{ijk}$:

$$\in^{lmn}=\in_{ijk} G^{il}G^{jm}G^{kn} = \sqrt{G}\left(e_{ijk}G^{il}G^{jm}G^{kn}\right) \tag{2.128}$$

It can be shown that the expression in brackets is the determinant of the matrix formed from the contravariant components of the metric tensor G^{ij} multiplied by e^{lmn}:

$$e_{ijk}G^{il}G^{jm}G^{kn} = e^{lmn}\det\begin{pmatrix} G^{il} & G^{im} & G^{in} \\ G^{jl} & G^{jm} & G^{jn} \\ G^{kl} & G^{km} & G^{kn} \end{pmatrix} \tag{2.129}$$

This is not a straightforward fact; a detailed derivation may be found in Flugge [91].

To calculate the determinant we recall that:

$$G_{ij}G^{jk} = \delta_i^k \tag{2.130}$$

That is, the matrixes formed by the covariant and contravariant components of the metric tensor are reciprocal (we have seen this before). Therefore, if we apply det to both sides of this equality:

$$\det(G_{ij}G^{jk}) = \det(\delta_i^k) = 1 \tag{2.131}$$

We know that the determinant of the product of two matrixes is the product of their determinants, therefore:

$$\begin{aligned} &\det(G_{ij}G^{jk}) = \det(G_{ij})\det(G^{jk}) = 1 \Rightarrow \\ &\det(G^{jk}) = \frac{1}{\det(G_{ij})} = \frac{1}{G} \end{aligned} \tag{2.132}$$

Now, returning to our expression for the contravariant components of the permutation tensor $\in^{lmn}$, we can see that:

$$\begin{aligned} &e_{ijk}G^{il}G^{jm}G^{kn} = e^{lmn}\frac{1}{G} \Rightarrow \\ &\in^{lmn} = \sqrt{G}\frac{1}{G}e^{lmn} = \frac{1}{\sqrt{G}}e^{lmn} \end{aligned} \tag{2.133}$$

There are three basic properties of the permutation tensor worth noting.

First of all, it is antimetric relative to any pair of its indices. That is:

$$\in^{ijk} = -\in^{jik}$$
$$\in^{ijk} = -\in^{kji} \qquad (2.134)$$
$$\ldots$$

The property of antimetry is opposite to the property of symmetry.

Secondly, if we take a symmetrical tensor $A_{ij} = A_{ji}$ and form a tensor product with the permutation tensor $\in^{ijk} A_{jk}$, it can be shown that this product is zero:

$$\in^{ijk} A_{jk} = 0 \qquad (2.135)$$

Indeed, for every i we have a double sum by j and k with 9 terms in total. Three of the terms of the sum contain $\in^{ijk}$ with $j = k$ and are therefore zero. The six remaining, those where $j \neq k$, can be grouped in pairs:

$$\left(\in^{i12} A_{12} + \in^{i21} A_{21}\right) + \left(\in^{i13} A_{13} + \in^{i31} A_{31}\right) + \left(\in^{i23} A_{23} + \in^{i32} A_{32}\right) \qquad (2.136)$$

Each grouped term in this sum is zero due to antimetry of the permutation tensor.

A similar relation holds for the covariant components of the permutation tensor and any symmetric tensor given by contravariant components:

$$A^{jk} = A^{kj} \Rightarrow \in_{ijk} A^{jk} = 0 \qquad (2.137)$$

The next section will describe one application of the permutation tensor.

2.6.7 Cross vector product

We have already used the cross (or vector) product of two vectors in our global Cartesian coordinate system both in vector form:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \qquad (2.138)$$

and in components:

$$\mathbf{a} = a^i \mathbf{k}_i$$
$$\mathbf{b} = b^i \mathbf{k}_i \qquad (2.139)$$

$$c_1 = a^2 b^3 - a^3 b^2$$

$$c_2 = -(a^1 b^3 - a^3 b^1)$$

$$c_3 = a^1 b^2 - a^2 b^1$$

Vector $\mathbf{c} = c^i \mathbf{k}_i$ is the cross product of vectors **a** and **b**, expressed via their coordinates.

From elementary geometry, we know that the cross product defined in this way has two important properties:

1. It is perpendicular to the plane where both **a** and **b** lie.
2. Its length is equal to the area dS of the parallelogram formed by the two vectors **a** and **b** (See Figure 2.14 below).

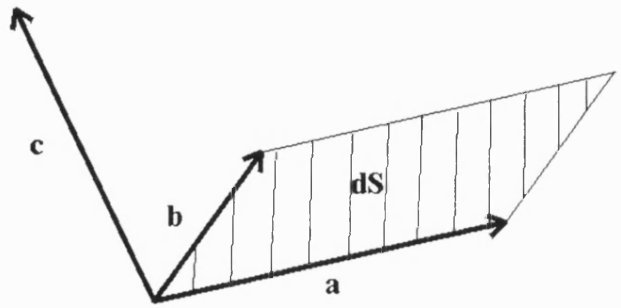


Figure 2.14: Cross product of two vectors

It is these two properties that make the notion of the cross product so valuable to the theory of elasticity (and the tensor analysis in general). The cross product provides an easy means of expressing an element of area.

Let us discuss this last point in more detail. Elements of area are small (generally infinitesimally small). They are normally used in expressions where the limiting case of the total area of the area element tending to zero is of importance. Taking this into account, it is clear that when dealing with infinitesimal area elements, their exact shape is not very important.

We can therefore represent such an element by a parallelogram (see Figure 2.14 above) with two vectors **a** and **b** defining its size. The cross product **c** of these vectors **a** and **b** will represent the area element: its length will be the area of the area element, and it will be perpendicular to the area element itself. From this we can see that we can represent the area

element by a vector **dS** with its length being equal to the area of the element and pointing along the normal (that is, perpendicular) to the area.

We will now derive a generic expression for the cross product of two vectors in any coordinate system.

Let us have two vectors **a** and **b**:

$$\begin{aligned} \mathbf{a} &= a^i \mathbf{G}_i \\ \mathbf{b} &= b^i \mathbf{G}_i \end{aligned} \tag{2.140}$$

We want to find the expression for the components of their cross product **c**:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \tag{2.141}$$

We have the expressions for the components of **c**, if $\mathbf{G}_i$ represent a Cartesian basis:

$$\begin{aligned} c_1 &= a^2 b^3 - a^3 b^2 \\ c_2 &= -(a^1 b^3 - a^3 b^1) \\ c_3 &= a^1 b^2 - a^2 b^1 \end{aligned} \tag{2.142}$$

We can rewrite these equalities in the following equivalent form using the permutation tensor in the Cartesian coordinate system:

$$\begin{aligned} c_1 &= e_{123}\left(a^2 b^3 - a^3 b^2\right) = e_{123} a^2 b^3 + e_{132} a^3 b^2 \\ c_2 &= e_{213} a^1 b^3 + e_{231} a^3 b^1 \\ c_3 &= e_{312} a^1 b^2 + e_{321} a^2 b^1 \end{aligned} \tag{2.143}$$

Using the summation convention, it is simply:

$$c_i = e_{ijk} a^j b^k \tag{2.144}$$

Or, recalling that in the Cartesian coordinate system, $e_{ijk} = \in_{ijk}$:

$$c_i = \in_{ijk} a^j b^k \tag{2.145}$$

Now, each of the terms on the right is a tensor, therefore c_i is a tensor as well, and it is the cross product of two vectors **a** and **b**, expressed in covariant components in any coordinate system.

Let us now find the expression for the cross product in contravariant components:

$$c^i = c_l G^{li} = \in_{ljk} G^{li} a^j b^k \tag{2.146}$$

Using the equality:

$$G_{ij} G^{jk} = \delta_i^k \tag{2.147}$$

We can continue the chain of transformations:

$$c^i = \in_{lmk} G^{li} \delta_j^m a^j b^k = \in_{lmk} G^{li} G^{mr} G_{rj} a^j b^k \tag{2.148}$$

and similarly:

$$c^i = \in_{lmn} G^{li} G^{mr} G_{rj} a^j \delta_k^n b^k = \in_{lmn} G^{li} G^{mr} G_{rj} a^j G^{ns} G_{sk} b^k \tag{2.149}$$

Or, regrouping the expression on the right:

$$c^i = \left(\in_{lmn} G^{li} G^{mr} G^{ns}\right)\left(a^j G_{rj}\right)\left(b^k G_{sk}\right) \tag{2.150}$$

Now we can see that each of the expressions in the brackets represents lowering or raising indices in the corresponding entities, and can be rewritten as follows:

$$c^i = \in^{irs} a_r b_s \tag{2.151}$$

Or, renaming the summation indices:

$$c^i = \in^{ijk} a_j b_k \tag{2.152}$$

This equality provides us with the expression for the contravariant components of the cross product of two vectors **a** and **b**.

2.6.8 Covariant derivative

We want to measure the rate of change of a vector $\mathbf{a} = a^i \mathbf{G}_i$, that is, its spatial derivative:

$$\mathbf{a}_{,j} = \frac{\partial}{\partial x^j}\mathbf{a} = \frac{\partial}{\partial x^j}\left(a^i \mathbf{G}_i\right) \tag{2.153}$$

If $\mathbf{G}_i$ represent the Cartesian coordinate system, then the vector derivative can be expressed in terms of the derivative of its components:

$$\mathbf{a}_{,j} = \frac{\partial}{\partial x^j}\left(a^i \mathbf{G}_i\right) = \frac{\partial a^i}{\partial x^j}\mathbf{G}_i = a^i_{,j}\mathbf{G}_i \tag{2.154}$$

We can do this because in the Cartesian coordinate system $\mathbf{G}_i$ are constant. Indeed, in the Cartesian coordinate system any change in a vector is due to the change of its coordinates.

In the general case, when $\mathbf{G}_i$ are not constant, it is no longer true. We can easily imagine an example where a vector with constant components is not constant (take the polar coordinate system for example: a vector with constant components will be rotating when moving along the angular coordinate x^2) and similarly a vector with changing components can actually be a constant vector.

Let us return to the general expression for a derivative of a vector, and expand it:

$$\mathbf{a}_{,j} = \frac{\partial}{\partial x^j}\left(a^i \mathbf{G}_i\right) = a^i_{,j}\mathbf{G}_i + a^i \mathbf{G}_{i,j} \tag{2.155}$$

Let us consider $\mathbf{G}_{i,j}$. It is a vector; we can therefore resolve it into components in the covariant or contravariant bases:

$$\mathbf{G}_{i,j} = \Gamma^k_{ij}\mathbf{G}_k = \Gamma_{ijk}\mathbf{G}^k \tag{2.156}$$

If we dot multiply both sides by first $\mathbf{G}^k$ and then by $\mathbf{G}_k$, we will get the following equalities for Γ^k_{ij} and Γ_{ijk} directly:

$$\begin{aligned} \mathbf{G}_{i,j} \cdot \mathbf{G}^k &= \Gamma^k_{ij} \\ \mathbf{G}_{i,j} \cdot \mathbf{G}_k &= \Gamma_{ijk} \end{aligned} \tag{2.157}$$

These two sets of symbols Γ^k_{ij} and Γ_{ijk} are called *Christoffel symbols*.

From their definition, we can note that:

$$\Gamma_{ij}^{k} G_{kl} = \mathbf{G}_{i,j} \cdot \mathbf{G}^{k} G_{kl} = \mathbf{G}_{i,j} \cdot \mathbf{G}_{l} = \Gamma_{ijl} \tag{2.158}$$

And similarly:

$$\Gamma_{ijk} G^{kl} = \Gamma_{ij}^{l} \tag{2.159}$$

We can see that the third index can be lowered and raised like the index of a tensor. It is not true however for the other two indices, and in general neither Γ_{ij}^{l} nor Γ_{iji} are components of a third order tensor (that is why they are called symbols).

From the fact that:

$$\mathbf{G}_{i,j} = \frac{\partial}{\partial x^{j}}\left(\frac{\partial \mathbf{R}}{\partial x^{i}}\right) = \frac{\partial^{2}\mathbf{R}}{\partial x^{j}\partial x^{i}} = \frac{\partial^{2}\mathbf{R}}{\partial x^{i}\partial x^{j}} = \mathbf{G}_{j,i} \tag{2.160}$$

And from their definition, we can conclude that the Christoffel symbols are symmetric with respect to the first two indices:

$$\begin{aligned} \Gamma_{ij}^{k} &= \Gamma_{ji}^{k} \\ \Gamma_{ijk} &= \Gamma_{jik} \end{aligned} \tag{2.161}$$

We will now return to the differentiation formulae for the vector **a**:

$$\mathbf{a}_{,j} = a_{,j}^{i}\mathbf{G}_{i} + a^{i}\mathbf{G}_{i,j} = a_{,j}^{i}\mathbf{G}_{i} + a^{i}\Gamma_{ij}^{k}\mathbf{G}_{k} \tag{2.162}$$

After renaming the indices i and k and regrouping the expression on the right, we will get the following equality:

$$\mathbf{a}_{,j} = \left(a_{,j}^{i} + a^{k}\Gamma_{kj}^{i}\right)\mathbf{G}_{i} \tag{2.163}$$

Definition (Covariant derivative for a^{i}) (2.17)

The quantity $a^{i}\big|_{j}$ given by the following equality:

$$a^{i}\big|_{j} = a_{,j}^{i} + a^{k}\Gamma_{kj}^{i} \tag{2.164}$$

is called the covariant derivative of the vector a^{i}.

Let us now derive the expression for the covariant derivative, involving the covariant vector components a_i.

As we know, the following is always true:

$$\mathbf{G}_k \cdot \mathbf{G}^i = \delta_k^i \tag{2.165}$$

If we now differentiate both sides by x^j, we will get the following:

$$\begin{gathered} \mathbf{G}_{k,j} \cdot \mathbf{G}^i + \mathbf{G}_k \cdot \mathbf{G}^i_{,j} = 0 \\ \mathbf{G}_k \cdot \mathbf{G}^i_{,j} = -\Gamma^i_{kj} \end{gathered} \tag{2.166}$$

The above equality is equivalent to the following one:

$$\mathbf{G}^i_{,j} = -\Gamma^i_{kj}\mathbf{G}^k \tag{2.167}$$

We will now consider the derivative of the vector expanded through the covariant components of the vector:

$$\mathbf{a}_{,j} = \frac{\partial}{\partial x^j}\left(a_i \mathbf{G}^i\right) = a_{i,j}\mathbf{G}^i + a_i \mathbf{G}^i_{,j} \tag{2.168}$$

Using the above-derived expression for $\mathbf{G}^i_{,j}$, we can re-write it as follows:

$$\mathbf{a}_{,j} = a_{i,j}\mathbf{G}^i - a_i\Gamma^i_{kj}\mathbf{G}^k \tag{2.169}$$

Or, after renaming indices and regrouping:

$$\mathbf{a}_{,j} = \left(a_{i,j} - a_k\Gamma^k_{ij}\right)\mathbf{G}^i \tag{2.170}$$

Definition (Covariant derivative for a_i) *(2.18)*

The quantity $a_i\big|_j$ given by the following equality:

$$a_i\big|_j = a_{i,j} - a_k\Gamma^k_{ij} \tag{2.171}$$

is the covariant derivative of the vector a_i.

Both expressions for the covariant derivative $a^i\big|_j$ and $a_i\big|_j$ (in contravariant and covariant vector components respectively) represent different forms of the same second order tensor. If this is true (as we will prove), it means we are in fact dealing with one entity - the second order vector derivative tensor. The two alternative expressions we have just derived are two different representations of this tensor. In total, as with any other second order tensor, this tensor has four different forms: $a_i\big|_j$, $a^i\big|_j$, $a_i\big|^j$ and $a^i\big|^j$.

Let us now show that indeed $a_i\big|_j$ is a second order tensor (in its covariant form). To do this, we need to show that it transforms as a second order tensor from one coordinate system into the other:

$$a'_{i'}\big|_{j'} = \beta^i_{i'}\beta^j_{j'}a_i\big|_j \tag{2.172}$$

Indeed, we have:

$$\mathbf{a} = a_i\mathbf{G}^i = a'_{i'}\mathbf{G}'^{i'} \tag{2.173}$$

We note that:

$$\mathbf{a}_{,j} = \beta^{j'}_j\mathbf{a}_{,j'} \tag{2.174}$$

Therefore, on the one hand:

$$\mathbf{a}_{,j} = a_i\big|_j\mathbf{G}^i \tag{2.175}$$

And on the other hand:

$$\mathbf{a}_{,j} = \beta^{j'}_j\mathbf{a}_{,j'} = \beta^{j'}_j a'_{i'}\big|_{j'}\mathbf{G}'^{i'} = \beta^{j'}_j a'_{i'}\big|_{j'}\beta^{i'}_i\mathbf{G}^i = \left(\beta^{j'}_j\beta^{i'}_i a'_{i'}\big|_{j'}\right)\mathbf{G}^i \tag{2.176}$$

Equating two expressions for $\mathbf{a}_{,j}$ we obtain the following:

$$a_i\big|_j\mathbf{G}^i = \left(\beta^{j'}_j\beta^{i'}_i a'_{i'}\big|_{j'}\right)\mathbf{G}^i \tag{2.177}$$

Multiplying both sides by $\mathbf{G}_k$ we finally obtain:

$$a_i\big|_j = \beta^{j'}_j\beta^{i'}_i a'_{i'}\big|_{j'} \tag{2.178}$$

This proves the tensor character of $a_i\big|_j$. Similarly, it can be proved that $a^i\big|_j$ is a tensor. Two remaining forms of the derivative tensor we have mentioned before $a_i\big|^j$ and $a^i\big|^j$ are obtained by raising the index.

One interesting point about the expression for the covariant derivative $a_i\big|_j$ (or any other form of the tensor):

$$a_i\big|_j = a_{i,j} - a_k\Gamma_{ij}^k \tag{2.179}$$

should be noted.

As we have just shown, $a_i\big|_j$ is a tensor and therefore the expression $a_{i,j} - a_k\Gamma_{ij}^k$ on the right is also a tensor. We know that $a_{i,j}$ is definitely not a tensor and neither is Γ_{ij}^k in respect to the indices i and j (and therefore $a_k\Gamma_{ij}^k$). But the remarkable thing is, the sum of these two non-tensors is itself a tensor. From this example and the facts about tensors we have introduced in the previous sections, we can make the following general conclusion: a product of two tensors and the sum of two tensors is a tensor; a product of a tensor and a non-tensor is not a tensor, and finally the sum of two non-tensors can be a tensor.

As we can now see, the expression for the derivative of a vector quantity expressed in the coordinate form involves the use of the Christoffel symbols. These quantities describe the material properties of the coordinate system and relate to the rate of change of the base vectors. We can see from the expression for the covariant derivative, that the Christoffel symbols take part in the correction term that compensates the variation in the basis of the underlying coordinate system to the usual partial derivative $a_{i,j}$.

As can easily be seen from the definition of the Christoffel symbols, in the case of the Cartesian coordinate system the Christoffel symbols are all zero. Therefore, in this case the covariant derivative coincides with the usual partial derivative (as one would expect).

We will now proceed to introduce the covariant derivative for the tensors of the orders other then 1. The formal derivation of the formulae is not material and the reader is referred to Flugge or any other text on tensor analysis.

If we have a scalar φ (tensor of zero-th order), its covariant derivative is the same as a normal partial derivative:

$$\varphi_{,i} = \varphi\big|_i \tag{2.180}$$

That is, $\varphi_{,i}$ is a vector (a tensor of the first order).

From the tensors of the higher order, we will only consider second order tensors. It can be shown that the following equalities describe the covariant derivative of a second order tensor in its various forms:

$$\begin{aligned}
A^{ij}\big|_k &= A^{ij}_{,k} + A^{lj}\Gamma^i_{kl} + A^{il}\Gamma^j_{kl} \\
A^{\cdot j}_{i\cdot}\big|_k &= A^{\cdot j}_{i\cdot,k} - A^{\cdot j}_{l\cdot}\Gamma^l_{ik} + A^{\cdot l}_{i\cdot}\Gamma^j_{kl} \\
A^{i\cdot}_{\cdot j}\big|_k &= A^{i\cdot}_{\cdot j,k} + A^{l\cdot}_{\cdot j}\Gamma^i_{kl} - A^{i\cdot}_{\cdot l}\Gamma^l_{kj} \\
A_{ij}\big|_k &= A_{ij,k} - A_{lj}\Gamma^l_{ik} - A_{il}\Gamma^l_{jk}
\end{aligned} \tag{2.181}$$

Again, we can see the Christoffel symbols taking part in the expressions, this time in two terms for each covariant derivative. From these equalities it is not difficult to see what the expressions for the covariant derivatives for the tensors of higher order will be. There will be as many terms involving the Christoffel symbols as the order of the tensor, the sign will be dependant on whether the index is covariant or contravariant and so on.

Because the covariant derivative is a tensor, we can raise its differentiation index as in any other tensor:

$$a_i\big|^j = a_i\big|_k G^{kj} \tag{2.182}$$

We can now consider the following expression:

$$a_i\big|^j_k = \left(a_i\big|^j\right)\big|_k \tag{2.183}$$

$a_i\big|^j_k$ is a tensor of the third order. We can now form the contraction:

$$a_i|_j^j \tag{2.184}$$

This is the tensor form of the Laplace operator ∇, which in the Cartesian coordinate system is given by the following more familiar equality:

$$\nabla = \sum_k \frac{\partial^2}{(\partial x^k)} \tag{2.185}$$

We have now shown how to differentiate tensors. We are almost ready to start the introduction of expression for the internal elastic force; the only two things left are integration and the Stokes Theorem linking surface integral with the volume one. We will need these when we consider the reactive action of the body on the small element (as we recall, it will be a surface integral summing up the force of the rest of the body on the element through its boundary) during the body deformation.

2.6.9 Divergence of a vector field

Let us have a vector field $\mathbf{v} = v^i \mathbf{G}_i$, defined for any point x^i.

If $\mathbf{G}_i$ represent the Cartesian coordinate system, the divergence of $\mathbf{v}$ is a scalar function defined as:

$$div(\mathbf{v}) = \frac{\partial v^1}{\partial x^1} + \frac{\partial v^2}{\partial x^2} + \frac{\partial v^3}{\partial x^3} = \sum_k \frac{\partial v^k}{\partial x^k} \tag{2.186}$$

To find the expression for the divergence in the case of a general curvilinear coordinate system $\mathbf{G}_i$, we proceed as follows. First, we re-write the above-given equality using our notation for the derivative:

$$div(\mathbf{v}) = v^1_{,1} + v^2_{,2} + v^3_{,3} = v^k_{,k} \tag{2.187}$$

We note that in the Cartesian coordinate system $v^k_{,k} = v^k|_k$, the expression for the divergence now takes the following form:

$$div(\mathbf{v}) = v^k|_k \tag{2.188}$$

In this tensor form, it is valid in any coordinate system. It can be noted that we could obtain an alternative expression for the divergence using the same arguments:

$$div(\mathbf{v}) = v_k|^k \tag{2.189}$$

Divergence has a simple physical meaning. If $\mathbf{v}$ is a velocity field of a body, its divergence will be related to the changes in the body's density, as it undergoes deformations. If for example it is known that $div(\mathbf{v}) > 0$ at a point, it can be shown that this will lead to the increase of the density at the point. Equally, if $div(\mathbf{v}) < 0$, it will lead to the decrease of the density. The condition $div(\mathbf{v}) = v^k|_k = 0$ is called the *incompressibility condition*; it simply states that the density of the body must remain unchanged during its deformation.

The following conclusion is true: $div(\mathbf{v})$ relates to the rate of change of the density (the exact law that binds the two together is called the continuity condition).

2.6.10 Integration and Gauss' divergence theorem

In section 2.5.1 we have derived the expression for the internal force term $\mathbf{E}(x^i,t)$, being the total force that the rest of the body exerts through the volume element's boundary.

It is clear that this term will be expressed via the integral over the volume element's boundary, namely of the form:

$$\mathbf{E}(x^i,t) = \int_{\partial B} \mathbf{E}^{surf}(\mathbf{dS}) \tag{2.190}$$

Where $\mathbf{dS}$ represents the boundary element, $\mathbf{E}^{surf}(\mathbf{dS})$ is the force transmitted through this boundary element (it represents the action of the rest of the body on this boundary element), and ∂B is the boundary of the volume element. Figure 2.15 provides a graphical illustration of the entities involved.

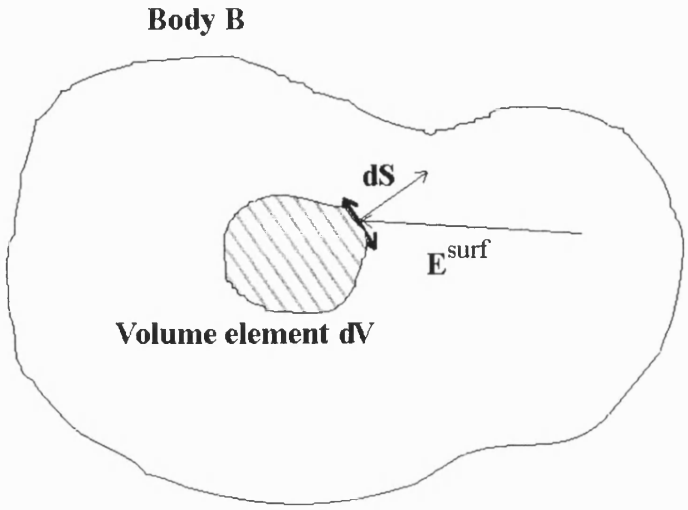


Figure 2.15: Force acting on an element of the boundary of the volume element dV

Gauss's divergence theorem provides us with an important tool for working with surface integrals of this kind. As will be shown later, this theorem will be used to obtain the final expression of the elastic force.

Let us suppose we have a volume V with its surface ∂V and a vector field $\mathbf{v}$. We choose an area element on the boundary ∂V, and denote it by $\mathbf{dS}$ (it is normal to the boundary and points outwards).

This surface element $\mathbf{dS}$ is small; we can therefore assume that the vector field is constant throughout the surface area element $\mathbf{dS}$ (we can measure it at any point x^i).

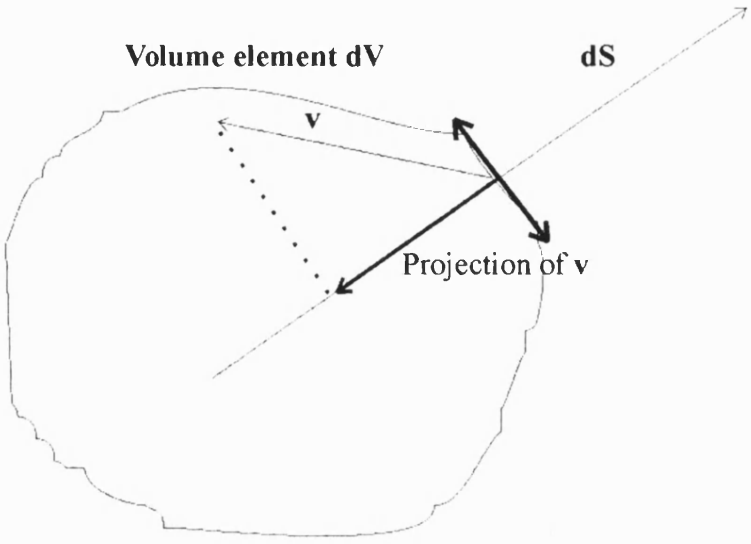


Figure 2.16: Volume element and vector field

We now want to calculate the projection of $\mathbf{v}$ onto $\mathbf{dS}$ (on the picture it is shown in bold):

$$proj(\mathbf{v}) = \mathbf{v} \cdot \mathbf{dS} \tag{2.191}$$

We now want to calculate the total of this projection over the whole boundary ∂V :

$$T = \int_{\partial V} proj(\mathbf{v}) = \int_{\partial V} \mathbf{v} \cdot \mathbf{dS} \tag{2.192}$$

T is called the flux of the vector field $\mathbf{v}$ through the boundary ∂V .

Gauss' divergence theorem states that this quantity T can be expressed through the volume integral of the divergence of the field $\mathbf{v}$ as follows:

$$T = \int_{\partial V} \mathbf{v} \cdot \mathbf{dS} = \int_{V} div(\mathbf{v}) \cdot dV \tag{2.193}$$

We will omit the proof of this theorem (the reader can refer to Flugge or any other text on elasticity), but rather give an example to help explain physical meaning.

Let us suppose V is a volume in the space through which a fluid is flowing, and $\mathbf{v} = \mathbf{v}(\mathbf{R},t)$ is the velocity of the fluid at the point in space $\mathbf{R}$ at the time t. If we now choose a boundary area element $\mathbf{dS}$ (normal to the boundary and pointing outwards as usual) and take its length to be one unit of area, then the expression:

$$proj(\mathbf{v}) = \mathbf{v} \cdot \mathbf{dS} \tag{2.194}$$

will represent the rate of flow of the fluid through this area element $\mathbf{dS}$ at an instant of time t. If it is positive, the fluid is flowing out, if it is negative, it is flowing in. If it is zero, the velocity $\mathbf{v}$ is zero and therefore the fluid is stationary around the element $\mathbf{dS}$.

We now consider the total flux of the fluid through the whole of the boundary ∂V :

$$T = \int_{\partial V} \mathbf{v} \cdot \mathbf{dS} \tag{2.195}$$

This represents the total flux of the fluid through the volume V . If it is positive, more fluid is leaving the volume than entering it, if it is negative, more fluid is entering the volume than leaving it. If it is zero however, we cannot simply conclude that the velocity $\mathbf{v}$ is zero everywhere on the boundary, because we have an integral (a sum) that is zero. However, we can say that if T is zero, the total amount of the fluid coming in is equal to the total amount coming out.

Let us consider the case when T is negative. In this case, more fluid is flowing in than flowing out. Because it is reasonable to assume that fluid does not just disappear inside our volume V, we arrive at the conclusion that its density inside the volume must increase. As we have noted before, $div(\mathbf{v})dV$ relates to the density, we can see there is a clear connection between $\mathbf{v}\cdot\mathbf{dS}$ (the rate of flow through the boundary) and $div(\mathbf{v})dV$ (the rate of change of the density of a volume element inside the volume V).

Gauss' theorem provide a precise description of this connection:

$$\int_{\partial V}\mathbf{v}\cdot\mathbf{dS}=\int_{V}div(\mathbf{v})\cdot dV \qquad (2.196)$$

2.7 The Internal Elastic Force

We have finished the preparatory considerations required for us to introduce and obtain formally the expression for the elastic (internal) force for a volume element.

The following sections will use the introduced entities and notions, and will lead to the formal derivation of the elastic force first for internal volume elements, and then for any volume elements.

2.7.1 Displacement field

So far, when considering the body we have always been dealing with the functional mapping **R**, giving the absolute positions of its points in our global Cartesian coordinate system K.

We have already obtained a motion law, involving **R**:

$$\ddot{\mathbf{R}}(x^i,t)\rho(x^i,t)=\mathbf{F}(x^i,t)+\mathbf{E}(x^i,t)=,\forall x^i\in I,\forall t>0 \qquad (2.197)$$

It is not convenient however to try to use **R** to obtain the formulae for $\mathbf{E}(x^i,t)$. Instead, it is more convenient to use the so-called displacement field, which we will formally introduce in this section.

Definition (displacement field) (2.19)

A vector function $\mathbf{U}(x^i,t)$ given by the following equality:

$$\mathbf{U}(x^i,t) = \mathbf{R}(x^i,t) - \mathbf{R}(x^i) = \mathbf{R}(x^i,t) - \mathbf{R}(x^i,0) \qquad (2.198)$$

is called a displacement field for the body B. It represents the path made by each individual point throughout our experiment (see the following Figure 2.17).

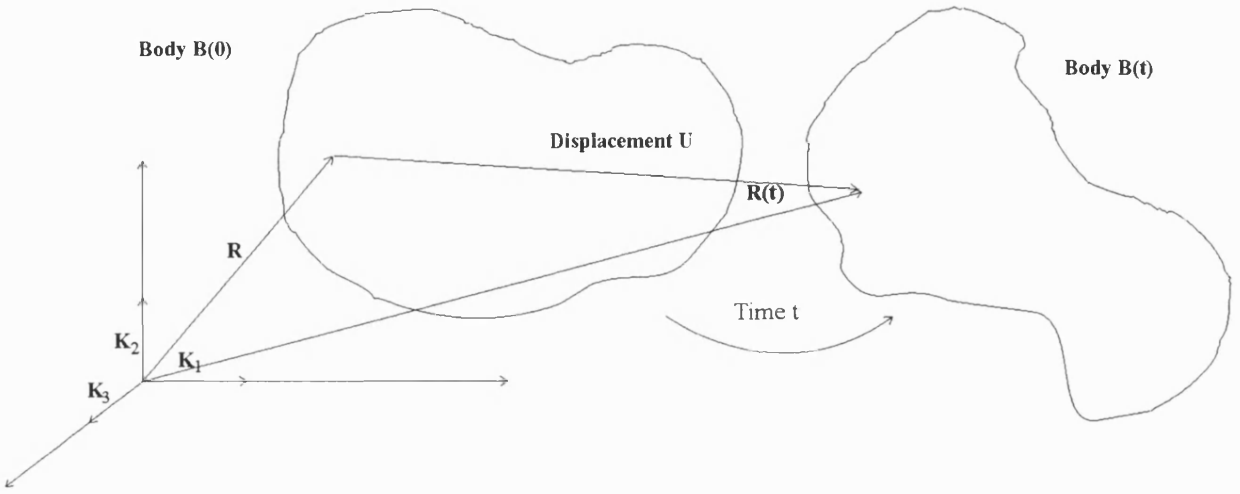


Figure 2.17: Displacement field U

This vector field is defined for every point on the body in its original undeformed state, which will call the *reference frame*. Therefore, the vector field **U** when given by components will be assumed to be resolved in the basis in the reference frame:

$$\mathbf{U}(x^i,t) = U^i(x^i,t)\mathbf{G}_i(x^i) = U^i(x^i,t)\mathbf{G}_i(x^i,0) \qquad (2.199)$$

The same will be true for all other quantities we will introduce in the following sections: strain and stress tensors, and, finally, the elastic force.

As we can see, the position vector $\mathbf{R}(x^i,t)$ and the displacement field $\mathbf{U}(x^i,t)$ are tied to each other in the relationship given above. Therefore, we can use either of them in any expression without limitation, with the other being uniquely dependent on it.

2.7.2 Strain tensor

We start with a basic measure of the deformation: the strain tensor.

We have two sets of base vectors: one on the body in its undeformed state $\mathbf{G}_i(x^i,0)$, and the other, in its deforming state $\mathbf{G}_i(x^i,t), t>0$. Similarly, we have the metric tensor of the undeformed body $G_{ij}(x^i,0)$, and the metric tensor of the deforming body $G_{ij}(x^i,t)$.

When we start the experiment, we have the body in its undeformed form; we have its base vectors $\mathbf{G}_i(x^i,0)$ and its metric tensor $G_{ij}(x^i,0)$. As the body starts to deform, the coordinate system we have attached to it also starts to deform into the one given by the base vectors $\mathbf{G}_i(x^i,t)$ and the metric tensor $G_{ij}(x^i,t)$.

We have seen in the examples that the metric tensor describes metric properties of the body. It is therefore natural to use the metric tensors of the undeformed and deformed bodies in the definition of the *strain* tensor: the kinematic measure of the deformation.

Definition (strain tensor) *(2.20)*

The quantity given by the following equality:

$$\gamma_{ij}(x^i,t)=G_{ij}(x^i,t)-G_{ij}(x^i,0) \qquad (2.200)$$

is called the strain tensor. It measures the pure deformation of the body (of the metric of the body to be more precise), and therefore serves as a measure of strain inside the body.

We will now show that it is actually a tensor. This is not an obvious statement because two metric tensors are expressed in different coordinate systems: $G_{ij}(x^i,t)$ in $\mathbf{G}_i(x^i,t)$ and $G_{ij}(x^i,0)$ in $\mathbf{G}_i(x^i,0)$.

We need to show that $\gamma_{ij}(x^i,t)$ show double covariant behaviour when transformed into another coordinate system. Namely:

$$\gamma'_{i'j'}(x'^{i'},t)=\beta^i_{i'}\beta^j_{j'}\gamma_{ij}(x^i,t)=\beta^i_{i'}\beta^j_{j'}\left(G_{ij}(x^i,t)-G_{ij}(x^i,0)\right) \qquad (2.201)$$

Let us look more carefully what we need to do.

At the beginning of our experiment ($t=0$), we have a body B and a coordinate system $\mathbf{G}_i(x^k)$ on it (given by a $\mathbf{R}(x^k)$). After the start of the experiment, the body begins to deform (becomes $\mathrm{B}(t)$), and all the quantities we have defined on it deform with it: base vectors deform into $\mathbf{G}_i(x^k,t)$, the metric tensor into $G_{ij}(x^k,t)$ and so on.

Here is what we need to establish: if we chose a *different* coordinate system $\mathbf{G}'_{i'}(x^{k'})$ (induced by a different mapping $\mathbf{R}'(x^{k'})$) associated with the body at the beginning of the experiment and worked with it throughout the experiment, what will the law of

transformation for all our quantities? We will now show that the transformation law will be the same at all times $t \geq 0$. More precisely, at the beginning of our experiment we have determined the transformation rule to be defined by a set of β:

$$\begin{aligned} \mathbf{G}'_{i'}\left(x^{k'}\right) &= \beta^{i}_{i'}\mathbf{G}_{i}\left(x^{k}\right) \\ \mathbf{G}'^{i'}\left(x^{k'}\right) &= \beta^{i'}_{i}\mathbf{G}^{i}\left(x^{k}\right) \end{aligned} \tag{2.202}$$

Then, as the experiment starts, the transformation rule will be:

$$\begin{aligned} \mathbf{G}'_{i'}\left(x^{k'},t\right) &= \beta^{i}_{i'}\mathbf{G}_{i}\left(x^{k},t\right) \\ \mathbf{G}'^{i'}\left(x^{k'},t\right) &= \beta^{i'}_{i}\mathbf{G}^{i}\left(x^{k},t\right) \end{aligned} \tag{2.203}$$

That is, the transformation coefficients $\beta^{i}_{i'}$ and $\beta^{i'}_{i}$ will remain unchanged, or, in other words, they do not depend on time.

Indeed, at the beginning of the experiment, we have two alternative mappings for our body B: $\mathbf{R}\left(x^{k}\right)$ and $\mathbf{R}'\left(x^{k'}\right)$. They create two alternative coordinate systems with base vectors $\mathbf{G}_{i}\left(x^{k}\right)$ and $\mathbf{G}'_{i'}\left(x^{k'}\right)$ respectively. We know that these two alternative mappings create a transformation between the coordinates:

$$\begin{aligned} x^{k'} &= x^{k'}\left(x^{i}\right) \\ x^{i} &= x^{i}\left(x^{k'}\right) \end{aligned} \tag{2.204}$$

We know how the two sets of base vectors are connected to each other:

$$\begin{aligned} \mathbf{G}'_{i'}\left(x^{k'}\right) &= \beta^{i}_{i'}\mathbf{G}_{i}\left(x^{k}\right) \\ \mathbf{G}'^{i'}\left(x^{k'}\right) &= \beta^{i'}_{i}\mathbf{G}^{i}\left(x^{k}\right) \end{aligned} \tag{2.205}$$

Where β is defined as:

$$\beta^{i}_{i'} = \frac{\partial x^{i}}{\partial x^{i'}} \tag{2.206}$$

$$\beta_i^{i'} = \frac{\partial x^{i'}}{\partial x^i}$$

We can see from these formulae that, because neither x^i nor $x^{i'}$ depend on time, $\beta_{i'}^{i}$ and $\beta_i^{i'}$ do not depend on time either.

Therefore, we have:

$$\begin{aligned} \mathbf{G}'_{i'}\left(x^{k'},t\right) &= \beta_{i'}^{i}\mathbf{G}_i\left(x^k,t\right) \\ \mathbf{G}'^{i'}\left(x^{k'},t\right) &= \beta_i^{i'}\mathbf{G}^i\left(x^k,t\right) \end{aligned} \tag{2.207}$$

and:

$$\begin{aligned} &\beta_{i'}^{j}\beta_{j'}^{j}\gamma_{ij}\left(x^i,t\right) = \beta_{i'}^{j}\beta_{j'}^{j}G_{ij}\left(x^i,t\right) - \beta_{i'}^{j}\beta_{j'}^{j}G_{ij}\left(x^i,0\right) = \\ &= G'_{i'j'}\left(x^{i'},t\right) - G'_{i'j'}\left(x^{i'},0\right) = \gamma'_{i'j'}\left(x'^{i'},t\right) \end{aligned} \tag{2.208}$$

This concludes our proof that indeed the strain tensor is a tensor in the coordinate system $\mathbf{G}_i\left(x^k\right)$ on the reference frame.

From the definition of the strain tensor, we can conclude that it is symmetric:

$$\gamma_{ij}\left(x^i,t\right) = \gamma_{ji}\left(x^i,t\right) \tag{2.209}$$

Let us return to the expression for the strain tensor:

$$\gamma_{ij}\left(x^i,t\right) = G_{ij}\left(x^i,t\right) - G_{ij}\left(x^i\right) \tag{2.210}$$

The strain tensor is expressed as a difference of the metric tensors at the current time t and the beginning of the experiment ($t = 0$), with the components expressed in the coordinate system $\mathbf{G}_i\left(x^k\right)$ associated with the reference frame. We also have the displacement field defined for every point on the body in the undeformed state and resolved in the coordinate system $\mathbf{G}_i\left(x^k\right)$. We can therefore expect that the strain tensor can be expressed as a function of the displacement field $\mathbf{U}\left(x^i,t\right)$.

Let us choose a point x^i and fix it. Then we choose a small displacement dx^i away from this point; $\mathbf{ds} = \mathbf{ds}(0)$ will be the projection of this displacement onto the undeformed body and $\mathbf{ds}(t)$, the projection of this displacement onto the deformed body.

We have the following equalities:

$$\mathbf{ds} = \mathbf{ds}(0) = \mathbf{G}_i dx^i$$
$$\mathbf{ds}(t) = \mathbf{G}_i(t)\, dx^i \tag{2.211}$$

Here $\mathbf{G}_i(t) \equiv \mathbf{G}_i(x^i, t)$ and $\mathbf{G}_i \equiv \mathbf{G}_i(x^i)$, we simply omit mentioning dependency on x^i.

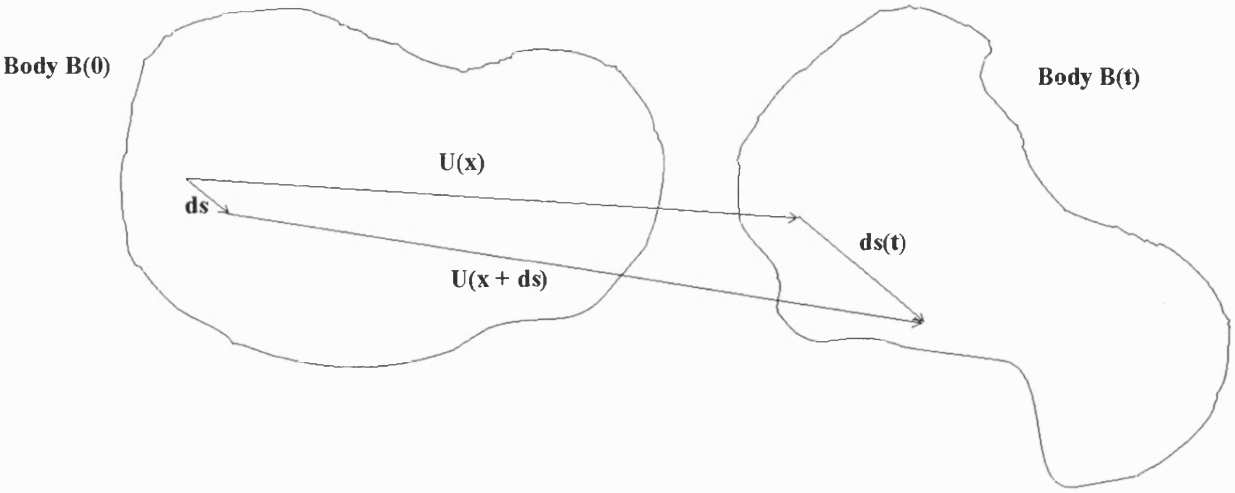


Figure 2.18: Deforming line element ds

Looking at Figure 2.18 we can write the following vector equation:

$$\mathbf{ds} + \mathbf{U}(x^i + dx^i) = \mathbf{U}(x^i) + \mathbf{ds}(t) \tag{2.212}$$

Because dx^i is small, we can expand the term $\mathbf{U}(x^i + dx^i)$:

$$\mathbf{U}(x^i + dx^i) = \mathbf{U}(x^i) + U^k\big|_l dx^l \mathbf{G}_k \tag{2.213}$$

Substituting this into the above-mentioned equation, we obtain the following vector equality:

$$\mathbf{ds}(t) = \mathbf{ds} + U^k\big|_l dx^l \mathbf{G}_k \tag{2.214}$$

We will now calculate the square of the length of this deformed line element $\mathbf{ds}(t)$:

$$ds(t)^2 = \mathbf{ds}(t) \cdot \mathbf{ds}(t) = \left(\mathbf{ds} + U^k\big|_l dx^l \mathbf{G}_k\right) \cdot \left(\mathbf{ds} + U^m\big|_n dx^n \mathbf{G}_m\right) \tag{2.215}$$

By the use of equivalent transformations we can transform this equality into the following form:

$$ds(t)^2 = \mathbf{ds}\cdot\mathbf{ds} + \mathbf{ds}\cdot\left(U^m\big|_n dx^n \mathbf{G}_m\right) + \mathbf{ds}\cdot\left(U^k\big|_l dx^l \mathbf{G}_k\right) + \left(U^k\big|_l dx^l \mathbf{G}_k\right)\cdot\left(U^m\big|_n dx^n \mathbf{G}_m\right) =$$

$$= ds^2 + \left(\mathbf{G}_i dx^i\right)\cdot\left(U^m\big|_n dx^n \mathbf{G}_m\right) + \left(\mathbf{G}_j dx^j\right)\cdot\left(U^k\big|_l dx^l \mathbf{G}_k\right) + \left(U^k\big|_l dx^l \mathbf{G}_k\right)\cdot\left(U^m\big|_n dx^n \mathbf{G}_m\right) =$$

$$= ds^2 + G_{im}U^m\big|_n dx^i dx^n + G_{jk}U^k\big|_l dx^l dx^j + G_{km}U^k\big|_l U^m\big|_n dx^n dx^l = \tag{2.216}$$

$$= ds^2 + U_i\big|_j dx^i dx^j + U_j\big|_i dx^i dx^j + U_m\big|_j U^m\big|_i dx^i dx^j$$

To obtain the last expression, we have renamed some summation indices, and used the metric tensor G_{im} to lower indices in the three terms.

To summarize, we have obtained the following equality:

$$ds(t)^2 = ds^2 + \left(U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j\right) dx^i dx^j \tag{2.217}$$

This can be further transformed, if we recall the expressions for $ds(t)^2$ and ds^2:

$$\left(G_{ij}(t) - G_{ij}\right) dx^i dx^j = \left(U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j\right) dx^i dx^j \tag{2.218}$$

Or:

$$\gamma_{ij} dx^i dx^j = \left(U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j\right) dx^i dx^j \tag{2.219}$$

The displacement dx^i is arbitrary. If we now choose it from the following set of values in turn: $\{dx^1 = 1, dx^2 = 0, dx^3 = 0\}, \ldots, \{dx^1 = 0, dx^2 = 0, dx^3 = 1\}$

We will obtain the final expression for the strain tensor:

$$\gamma_{ij} = U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j \tag{2.220}$$

This equality provides us with the expression of the strain tensor in terms of the displacement field $\mathbf{U}(x^i, t)$.

This equality is correct for any displacement field $\mathbf{U}(x^i, t)$. However, it is not easy to handle because of its non-linear term $U^m\big|_i U_m\big|_j$.

Clearly, the only way to simplify the full expression for the strain tensor is to drop this non-linear term altogether. To make it possible, we have to make an assumption that will justify us doing it. The assumption is that the first derivative of our displacement field is globally small. That is, for any point and at any time, the length of the derivative $\mathbf{U}_{,j}(x^i,t)$ of the displacement vector $\mathbf{U}(x^i,t)$ is smaller than a given small value $\varepsilon > 0$:

$$\left|\mathbf{U}_{,j}(x^i,t)\right| \le \varepsilon, \forall x^i, \forall t \ge 0, 1 \le j \le 3 \tag{2.221}$$

With such an assumption in hand, we can drop the non-linear term and obtain the following simple formula for the strain tensor:

$$\gamma_{ij} = U_i\big|_j + U_j\big|_i \tag{2.222}$$

As we already mentioned, the strain tensor relates to the geometrical deformation of the body; no forces are mentioned yet. This relationship provides the *kinematic* description of the body deformation.

In engineering, instead of the above formulae for the strain, it is more common to use another, slightly different form:

$$\varepsilon_{ij} = \frac{1}{2}\gamma_{ij} \tag{2.223}$$

Which leads to two new forms for the strain tensor:

$$\begin{aligned} \varepsilon_{ij} &= \frac{1}{2}\left(U_i\big|_j + U_j\big|_i\right) \\ \varepsilon_{ij} &= \frac{1}{2}\left(U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j\right) \end{aligned} \tag{2.224}$$

We will use this engineering strain ε_{ij} instead of γ_{ij} as it is more frequently used in applications.

We will now proceed to introduce the stress tensor, which relates to the distributed forces inside the body.

2.7.3 Stress tensor

Our aim in this section is to introduce a quantity that will relate to internal forces in the deforming body. As we can see from (2.28), the expression for the internal elastic force involves a surface integral summing forces acting through the boundary of a volume element. More precisely, the integral summing the forces transmitted through the elements of area over the whole boundary of the volume element.

We have shown in section 2.6.7 that the easiest way to represent an area element is by a vector, where the length of the vector is the area of the area element and the direction is perpendicular to the element itself.

We will proceed in three stages:

(1) We will show that any element of area (represented by a vector) can be presented as a sum of elements of area of three sides of a pyramid constructed in a special way.

(2) We will show that for these pyramids, any force acting on one pyramid's face can be resolved into a sum of forces acting on the other three faces.

(3) We will introduce the stress tensor itself.

Assume we have an area element represented by a vector $\mathbf{dA}$. As we have discussed before, the vector points in the perpendicular direction to the plane in which the area element lies, and its length is the area of the area element. The exact shape of the area is of no importance to us (because, as we have discussed before, we will be dealing with infinitesimal area elements).

As we know, we can represent $\mathbf{dA}$ as a vector product of two vectors:

$$\mathbf{dA} = \mathbf{s} \times \mathbf{r} \qquad (2.225)$$

Because the shape of the area is not important, we have few restrictions on the two vectors $\mathbf{s}$ and $\mathbf{r}$, as long as the equality $\mathbf{dA} = \mathbf{s} \times \mathbf{r}$ holds.

The length of $\mathbf{dA}$ is the area of the parallelogram based on $\mathbf{s}$ and $\mathbf{r}$, therefore, the length of $\frac{1}{2}\mathbf{dA}$ is the area of the triangle ABC based on $\mathbf{s}$ and $\mathbf{r}$ (see Figure 2.19 below).

We now proceed as follows: we choose a point O near the triangle so that the lines passing through O in the directions of the covariant base vectors pass through our points A , B and C (see Figure 2.19 below).

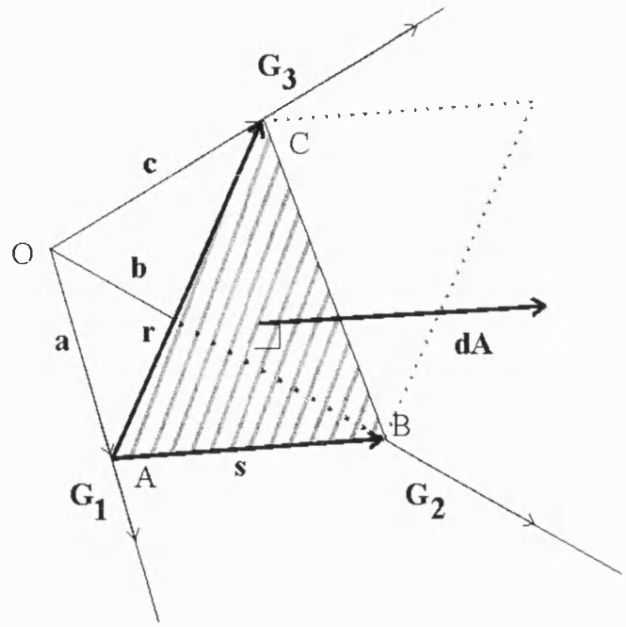


Figure 2.19: Area element and associated volume pyramid

Let us now show this point O always exists; this is quite an informal proof. Indeed, for every point near the plane we can find at least three points in which the lines, passing through it in the directions of the covariant base vectors, intersect our plane (where the area element is). For those three points we can calculate the area of the triangle and compare with the one we need to find: $\frac{1}{2}|\mathbf{dA}|$. If we position O on the plane, we can clearly see that this area of the triangle is zero. As we move the point away from the plane, we can expect the area to grow continually. At certain point we can expect it to reach the desired $\frac{1}{2}|\mathbf{dA}|$. It is obviously formally possible that, as we move the point O away from the plane, the area of the triangle will never reach $\frac{1}{2}|\mathbf{dA}|$, where $\mathbf{dA}$ is an arbitrary vector. However, as we have mentioned before, we will be dealing with infinitesimal area elements, those that are very small (tending to zero). In this case it is enough for the triangle area to grow to any non-zero value as we move O away from the plane. Due to continuity it will assume any value in between, and it will be enough to be applicable to our infinitesimal element $\mathbf{dA}$.

We will now proceed to show that the area element for the triangle ABC ($\frac{1}{2}\mathbf{dA}$) can be presented as a sum of the area elements of the triangles OAB, OBC and OCA. Indeed:

$$\frac{1}{2}\mathbf{dA} = \frac{1}{2}\mathbf{r}\times\mathbf{s} = \frac{1}{2}(\mathbf{b}-\mathbf{a})\times(\mathbf{c}-\mathbf{a}) = \frac{1}{2}\mathbf{b}\times\mathbf{c} + \frac{1}{2}\mathbf{c}\times\mathbf{a} + \frac{1}{2}\mathbf{a}\times\mathbf{b} \qquad (2.226)$$

The three terms on the right are the area elements for the specified triangles, pointing inward to form the pyramid. Also, due to the way we constructed our pyramid we can note that for each term we have the following information:

- It is parallel to one of the contravariant base vectors: $\frac{1}{2}\mathbf{b}\times\mathbf{c}$ is parallel to $\mathbf{G}_1$, $\frac{1}{2}\mathbf{c}\times\mathbf{a}$ is parallel to $\mathbf{G}_2$ and $\frac{1}{2}\mathbf{a}\times\mathbf{b}$ is parallel to $\mathbf{G}_3$.
- Its length is the area of the corresponding triangle: $\frac{1}{2}|\mathbf{b}\times\mathbf{c}|$ is equal to the area of OBC, $\frac{1}{2}|\mathbf{c}\times\mathbf{a}|$ is equal to the area of OCA and $\frac{1}{2}|\mathbf{a}\times\mathbf{b}|$ is equal to the area of OAB.

We will now resolve the vector **dA** into covariant components:

$$\mathbf{dA} = dA_i\mathbf{G}^i \qquad (2.227)$$

Using the above considerations, we can see that:

$$\begin{aligned}
\frac{1}{2}\mathbf{b}\times\mathbf{c} &= \frac{1}{2}dA_1\mathbf{G}^1 \\
\frac{1}{2}\mathbf{c}\times\mathbf{a} &= \frac{1}{2}dA_2\mathbf{G}^2 \qquad (2.228) \\
\frac{1}{2}\mathbf{a}\times\mathbf{b} &= \frac{1}{2}dA_3\mathbf{G}^3
\end{aligned}$$

It should be noted that it would be wrong to conclude from these equalities that simply, for example, $dA_1 = |\mathbf{b}\times\mathbf{c}|$. It is wrong because, in general, $\mathbf{G}^1$ is not of unit length; the correct expression is: $dA_1 = \frac{|\mathbf{b}\times\mathbf{c}|}{|\mathbf{G}^1|}$. Therefore, in general dA_1 is *not equal* to the area of the element

represented by $\mathbf{b}\times\mathbf{c}$, but rather is *proportional* to it. Similar statements are valid for the other covariant components of $\mathbf{dA}$.

This completes the first stage, and we will now move on to forces.

We assume now that the pyramid is cut inside our body, and there are net forces acting on each of its four sides (see Figure 2.20 below).

The pyramid is in a state of equilibrium; therefore, the sum of all forces acting on each side must be zero:

$$\mathbf{dF}+\mathbf{dP}+\mathbf{dQ}+\mathbf{dR}=0$$
$$\mathbf{dF}=-\mathbf{dP}-\mathbf{dQ}-\mathbf{dR} \qquad (2.229)$$

We can see from the last equality that $\mathbf{dF}$ can be presented as a sum of three forces $-\mathbf{dP}$, $-\mathbf{dQ}$ and $-\mathbf{dR}$, acting on the corresponding *internal* sides of the pyramid (as opposed to $\mathbf{dP}$, $\mathbf{dQ}$ and $\mathbf{dR}$ acting on the *external* sides of the pyramid).

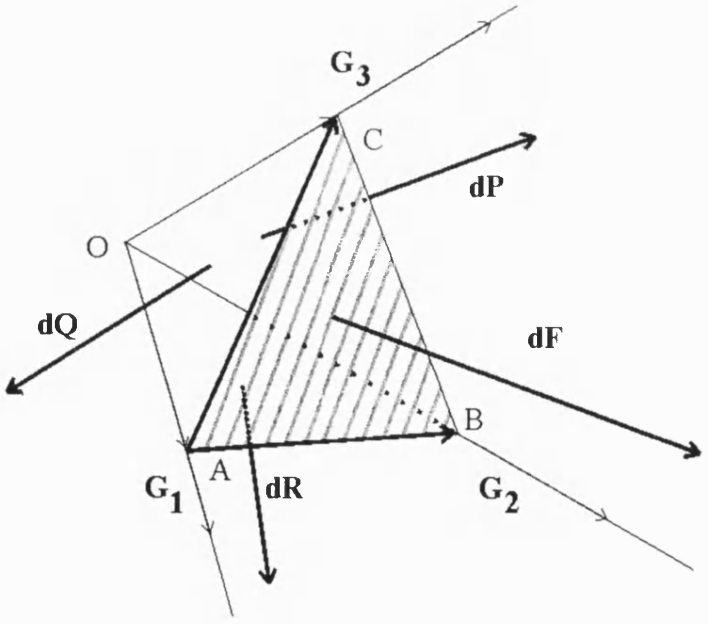


Figure 2.20: Forces acting on the sides of a pyramid inside the body

This concludes the second stage of our introduction of stress, and we are now ready to proceed with the introduction of the stress tensor.

We remember that $\mathbf{dA}$ represents the area element on which the force $\mathbf{dF}$ is acting.

Let us consider the force $-\mathbf{dP}$. It acts on the internal side of the pyramid, which is represented, as we have shown before, by an area element $\frac{1}{2}\mathbf{b}\times\mathbf{c}=\frac{1}{2}dA_1\mathbf{G}^1$. We will now resolve this force vector $-\mathbf{dP}$ into contravariant components, remembering that, because the force acts on the area element, it must be proportional to the area of the element (which is represented by dA_1):

$$-\mathbf{dP}=\sigma^{1j}dA_1\mathbf{G}_j \tag{2.230}$$

And similarly for the other two forces:

$$\begin{aligned}-\mathbf{dQ}&=\sigma^{2j}dA_2\mathbf{G}_j\\ -\mathbf{dR}&=\sigma^{3j}dA_3\mathbf{G}_j\end{aligned} \tag{2.231}$$

Recalling the expression for $\mathbf{dF}$, we can write the following equality:

$$\mathbf{dF}=\sigma^{1j}dA_1\mathbf{G}_j+\sigma^{2j}dA_2\mathbf{G}_j+\sigma^{3j}dA_3\mathbf{G}_j=\sigma^{ij}dA_i\mathbf{G}_j \tag{2.232}$$

Or, in terms of the components of $\mathbf{dF}$:

$$dF^j=\sigma^{ij}dA_i \tag{2.233}$$

Both dF^j and dA_i are components of vectors, therefore, the quantity σ^{ij} is a second order tensor, given in contravariant form. This tensor is called the *stress tensor*.

The stress tensor depends on the independent coordinates $\sigma^{ij}=\sigma^{ij}\left(x^k\right)$, and O is the point at which the value of the stress tensor is taken.

Now we can summarise what we have obtained. For any small area element inside our body, represented by a vector $\mathbf{dA}$, there exist a point O near the area element such that any internal force $\mathbf{dF}$ acting on the area element $\mathbf{dA}$ can be represented in the form $dF^j=\sigma^{ij}dA_i$, where σ^{ij} is the stress tensor taken at the point O.

It can be shown, by considering the equilibrium of an internal skew block of the material (see for example Flugge [91]), that the stress tensor is *symmetric*:

$$\sigma^{ij}=\sigma^{ji} \tag{2.234}$$

We can see now why it is impossible to represent internal (elastic) forces with one single vector; a symmetric second order tensor is required (that is, a quantity with six independent components rather than three for a vector).

2.7.4 Constitutive equations

In this section we will introduce the expressions, describing the relationship between the stress tensor and the strain tensor. The stress-strain relation represents internal *material* properties of the object we are modelling, and, therefore, the stress-strain relations are generally referred to as the constitutive equations.

We know the internal deformation is represented by the strain tensor, which measures the geometrical deformation of the body away from its original shape, regardless of the applied forces. On the other hand, the stress tensor is related to the internal forces, transmitted through internal area elements, regardless of how the body is deformed.

We will now find the relationship between the strain and the stress tensor. It is clear that one exists: if we apply forces to an elastic body it will deform, and, if we deform the body, it will result in an increase of the internal forces, trying to restore the original undeformed shape.

To proceed we will make the following assumption: for perfectly elastic bodies the deformation is linearly proportional to the applied force. This is a reasonable assumption - we know from the practical experience that a spring (the simplest example of an elastic body) conforms to this rule for small deformations.

If we make this assumption, it restricts the range of the applicability of the theory. However, there are several reasons why such an assumption is acceptable. First of all, we have already made an assumption for globally small gradients of the displacements (see section 2.7.2); secondly, most real elastic materials conform to this assumption of the linear relationship between the deformation and the elastic forces for small displacements.

Since we are dealing with second order tensors, the most generic form of a linear relationship between the stress and the strain tensors can be written as follows:

$$\sigma^{ij} = E^{ijlm}\varepsilon_{lm} \tag{2.235}$$

This equation is *Hooke's law* for a general solid. The quantity E^{ijlm} is a fourth order tensor, called the *elastic moduli.*

This stress-strain equation is a general form of the linear stress-strain relation where each individual strain component can exert (linear) influence over each individual stress component.

The tensor E^{ijlm} has 81 components in total; however, not all of them are independent. We know that both stress and strain tensors are symmetric. Therefore, E^{ijlm} is also symmetric in its first and second indices, and the third and fourth:

$$E^{ijlm} = E^{jilm}$$
$$E^{ijlm} = E^{ijml} \tag{2.236}$$

This reduces the number of independent components of E^{ijlm} to 36. This number can be further reduced to 21 after considering the strain energy (see Flugge [91]).

We, therefore, can finally conclude that in its general form, the tensor E^{ijlm} has 21 independent components.

In this general form, E^{ijlm} describes the stress-strain relation for an anisotropic body. That is, the body that can have different material properties in different directions (like crystals).

Most of the elastic solids, however, are isotropic: they have the same material properties regardless of the direction. In this special, but most practically applicable case, it can be shown that E^{ijlm} has only two independent components.

There are two pairs of the independent parameters that are used in the case of isotropic solids in the theory of elasticity. The first pair consists of the modulus of elasticity E and the Poisson ratio ν; the second pair consists of the two Lame's constants: λ and μ. The two pairs are obviously connected to each other, which we will show below.

The modulus of elasticity E and the Poisson ratio ν are most commonly used in engineering; they are most frequently used in applications. The formulae for the E^{ijlm} expressed in these parameters are given below:

$$E^{1111} = E^{2222} = E^{3333} = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \tag{2.237}$$

$$E^{1122} = E^{1133} = E^{2211} = E^{2233} = E^{3311} = E^{3322} = \frac{E\nu}{(1+\nu)(1-2\nu)}$$

$$E^{1212} = E^{1221} = E^{2112} = E^{2121} = \ldots = \frac{E}{2(1+\nu)}$$

$$E^{1223} = E^{1112} = E^{1123} = E^{2131} = \ldots = 0$$

Physically, E relates to the reaction of the material to stretching and squashing, and the Poisson ration relates to the bulging of the material when squashed, and narrowing in the middle when stretched.

The Lame's elastic constants are more convenient in theoretical manipulation since they afford easy tensor representation. The formulae for the E^{ijlm} expressed in the Lame's elastic constants are given below:

$$E^{ijlm} = \lambda G^{ij} G^{lm} + \mu\left(G^{il} G^{jm} + G^{im} G^{jl}\right) \tag{2.238}$$

The Lame's elastic constants have a simple relation, expressing them in terms of the modulus of elasticity and the Poisson ration:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$$

$$\mu = \frac{E}{2(1+\nu)} \tag{2.239}$$

Which formula for E^{ijlm} to use depends on the specific modelling task. If the solid possesses anisotropic properties, an appropriate expression for the elastic moduli will need to be obtained. In most cases, however, the solids are isotropic and the above-given expressions involving the elastic modulus and Poisson ratio, or the Lame's constants are sufficient.

In this thesis, when deriving the expression for the elastic force, we will use (2.239), expressing the elastic moduli E^{ijlm} through the Lame's constants.

2.8 Fundamental Equations of the Classical Theory of Elasticity

We have by now introduced all the necessary quantities and instruments in order to derive the expression for the elastic force for various volume elements and, later, the final forms of the equations of motion for a solid.

2.8.1 Equations of motion for internal points

In this section we will deal only with purely internal volume elements. We will firstly derive the expression for the internal elastic force for internal volume elements, and then give the final equations of motion for internal points of solids.

Let us recall the general expression, we have obtained at the beginning of this chapter, for the elastic force for a volume element:

$$\mathbf{E}(x^i,t) = \lim_{dV\to 0} \frac{1}{dV} \mathbf{E}_{total}(\Omega) \tag{2.240}$$

Here Ω is the internal volume element, dV is its volume and $\mathbf{E}_{total}(\Omega)$ is the total force the rest of the body exerts on the volume element through the element's boundary $\partial\Omega$.

We can now write down the expression for $\mathbf{E}_{total}(\Omega)$ in terms of the stress tensor. Indeed, we know that for any area element **dA** inside our solid, the force **F** transmitted through this element is:

$$F^i = \sigma^{ij} dA_j \tag{2.241}$$

As always, **dA** is the vector perpendicular to the plane where the area element lies and its length is equal to the size of the area.

Now the expression for the total force $\mathbf{E}_{total}(\Omega)$, which the rest of the solid exerts on the volume element through its boundary, can be obtained:

$$E^i_{total}(\Omega) = \int_{\partial\Omega} \sigma^{ij} dA_j \tag{2.242}$$

Here dA_j is the infinitesimal surface area element for our volume element. We cover the boundary of our volume element with these infinitesimal area elements and then sum the forces transmitted through those area elements using the integral over the whole boundary of the volume element.

For every fixed i, our integral sums the scalar products of the stress tensor (with the fixed index i) and the infinitesimal surface area elements dA_j. We can, therefore, use the Gauss divergence theorem:

$$E^i_{total}(\Omega) = \int_{\partial\Omega} \sigma^{ij} dA_j = \int_{\Omega} \sigma^{ij}\big|_j \, dY \tag{2.243}$$

Here dY is the volume of the infinitesimal volume elements, for which we calculate $\sigma^{ij}\big|_j$ and for which we then calculate the volume integral.

We now shrink our volume element Ω and make it infinitesimally small. In this case, we can say that our volume integral is:

$$\int_{\Omega} \sigma^{ij}\big|_j \, dY \cong \sigma^{ij}\big|_j \int_{\Omega} dY = \sigma^{ij}\big|_j \, dV \tag{2.244}$$

Using the full expression for the elastic force we can see that:

$$E^i = \lim_{dV \to 0} \frac{1}{dV} E^i{}_{total}(\Omega) = \lim_{dV \to 0} \frac{1}{dV} \left(\sigma^{ij}\big|_j \, dV \right) = \sigma^{ij}\big|_j \tag{2.245}$$

The elastic force E^i here is expressed in contravariant components in our curvilinear coordinate system $\mathbf{G}_i\left(x^k\right)$ (which is the curvilinear coordinate system on the reference frame):

$$\mathbf{E}\left(x^k, t\right) = E^i\left(x^k, t\right)\mathbf{G}_i = \sigma^{ij}\left(x^k, t\right)\big|_j \, \mathbf{G}_i \tag{2.246}$$

Now, after we have obtained the expression for the elastic force for internal volume elements (or for purely internal points), we can write down the equations of motion:

$$\ddot{\mathbf{R}}(x^i, t)\rho(x^i, t) = \mathbf{F}(x^i, t) + \mathbf{E}(x^i, t) = \mathbf{F}(x^i, t) + \sigma^{ij}\left(x^k, t\right)\big|_j \, \mathbf{G}_i\left(x^k\right) \tag{2.247}$$

This equation of motion is valid for any internal point and at any time.

Let us now transform this equation in its general form into the equation involving the displacement field. We remember that stress is resolved through strain via the constitutive equations (2.235) and the strain tensor itself is expressed via the displacement field (2.222).

We have:

$$\sigma^{ij} = E^{ijlm}\varepsilon_{lm} = \left(\lambda G^{ij} G^{lm} + \mu\left(G^{il} G^{jm} + G^{im} G^{jl}\right)\right)\varepsilon_{lm} \tag{2.248}$$

We now transfer ε_{lm} into the brackets, and use the fact that each G^{jm} raises one of the strain tensor's indices:

$$\sigma^{ij} = \lambda G^{ij} G^{lm} \varepsilon_{lm} + \mu \left(G^{il} G^{jm} \varepsilon_{lm} + G^{im} G^{jl} \varepsilon_{lm} \right) =$$
$$= \lambda G^{ij} G^{lm} \varepsilon_{lm} + \mu \left(\varepsilon^{ij} + \varepsilon^{ji} \right) = \lambda G^{ij} G^{lm} \varepsilon_{lm} + 2\mu \varepsilon^{ij} \tag{2.249}$$

To transform $G^{lm}\varepsilon_{lm}$ we proceed as follows. We can introduce formally a new tensor:

$$T^{lm}_{rs} = G^{lm} \varepsilon_{rs} \tag{2.250}$$

We can see that:

$$G^{lm} \varepsilon_{lm} = T^{lm}_{lm} \tag{2.251}$$

That is, T^{lm}_{lm} is the double contraction of the tensor T^{lm}_{rs}:

$$T^{lm}_{rm} = G^{lm} \varepsilon_{rm} = \varepsilon^{l}_{r}$$
$$T^{lm}_{lm} = \varepsilon^{l}_{l} \tag{2.252}$$

We therefore obtain the following expression for the stress tensor:

$$\sigma^{ij} = \lambda G^{ij} \varepsilon^{l}_{l} + 2\mu \varepsilon^{ij} \tag{2.253}$$

Recalling the expression for the strain, we can obtain the following set of equalities:

$$\varepsilon_{ij} = \frac{1}{2} \left(U_i \big|_j + U_j \big|_i \right)$$
$$\varepsilon^{i}_{i} = \frac{1}{2} \left(U^i \big|_i + U_i \big|^i \right) = U^i \big|_i \tag{2.254}$$
$$\varepsilon^{ij} = \frac{1}{2} \left(U^i \big|^j + U^j \big|^i \right)$$

We can now write the stress tensor in terms of the displacement field:

$$\sigma^{ij} = \lambda G^{ij} U^l \big|_l + \mu \left(U^i \big|^j + U^j \big|^i \right) \tag{2.255}$$

Finally, we can obtain the formulae for the derivative of the stress tensor:

$$\sigma^{ij} \big|_j = \left(\lambda G^{ij} U^l \big|_l + \mu \left(U^i \big|^j + U^j \big|^i \right) \right) \Big|_j = \tag{2.256}$$

$$= \lambda G^{ij} U^l\big|_{lj} + \mu\left(U^i\big|^j_j + U^j\big|^i_j\right) =$$

$$= \lambda U^l\big|^i_l + \mu\left(U^i\big|^j_j + U^j\big|^i_j\right)$$

If we now rename the indices, we will obtain the following equation:

$$\sigma^{ij}\big|_j = \lambda U^j\big|^i_j + \mu\left(U^i\big|^j_j + U^j\big|^i_j\right)$$

$$\sigma^{ij}\big|_j = (\lambda+\mu)U^j\big|^i_j + \mu U^i\big|^j_j \qquad (2.257)$$

This last equation we can use to obtain the final form of the equation of motion for the internal points of our (isotropic) solid:

$$\ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \left((\lambda+\mu)U^j\big|^i_j + \mu U^i\big|^j_j\right)\mathbf{G}_i\left(x^k\right) \qquad (2.258)$$

This equation is a vector equation for the unknown vector function $\mathbf{R}(x^i,t), t>0$, which represents the positions of the points of the solid during the experiment.

This equation of motion provides the solution to only a part of our original modelling problem. It only describes the trajectories of the strictly internal points, that is, it is valid in the situation when the boundary is constrained not to move for the duration of the experiment.

We will now obtain the equation of motion for the boundary points. This will obviously have to include the boundary forces.

2.8.2 Boundary conditions

At the moment we only have two pieces of information about the boundary points: we have the external force acting on each point and we know the elastic force, transmitted through any area element on the boundary.

More precisely, let us have a boundary point and a small boundary area element, represented by the vector $\mathbf{dA}$, $dA = |\mathbf{dA}|$. Then the external force acting on this area element is:

$$S^i dA \tag{2.259}$$

On the other hand, we can calculate the elastic force transmitted through this area element due to the action of the solid on this boundary element:

$$\sigma^{ij} dA_j \tag{2.260}$$

These two forces, the internal and external, must balance each other out, that is, the following equality must hold:

$$S^i dA = \sigma^{ij} dA_j \tag{2.261}$$

We can note that:

$$dA_j = n_j dA \tag{2.262}$$

Where n_j is the external normal to the boundary. We can now re-write the *equilibrium boundary conditions* in the following form:

$$S^i = \sigma^{ij} n_j \tag{2.263}$$

This is a set of static conditions that all the boundary points must satisfy at all times. It is static because it does not involve a derivative in time.

2.8.3 The full set of equations of motion

We can now combine the boundary conditions with the equations of motion for the internal points to obtain the full set of equations for all points in the solid:

$$\begin{cases} \ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \sigma^{ij}\left(x^k,t\right)\Big|_j \mathbf{G}_i\left(x^k\right), x^i \in \operatorname{int} I \\ S^i(x^i,t)\mathbf{G}_i\left(x^k\right) = \sigma^{ij}(x^i,t) n_j(x^i,t)\mathbf{G}_i\left(x^k\right), x^i \in \partial I \end{cases} \tag{2.264}$$

Practically, this system of equations means that we need to solve the dynamic equations for all the internal points and find the positions of the boundary points to satisfy the boundary conditions.

One can see from this system of equations how the external boundary forces influence the motion of the solid. From the definition of the stress tensor, we can see that σ^{ij} is calculated using the position of the point itself and its "immediate" neighbour for every direction (it is

ultimately the first spatial derivative of the displacement field), while $\sigma^{ij}\big|_j$ is calculated using two immediate neighbours of the point for each direction (because it is ultimately the second spatial derivative of the displacement field). Therefore, σ^{ij} on the boundary will be calculated using the boundary point and an adjacent internal point (as well as neighbouring points on the boundary of course, but this is not very important).

Let us now suppose the external boundary force at a certain boundary point was zero and then started to increase (the solid got into contact with another solid, for example). In this case, through the boundary condition, the external force will demand a corresponding increase in the stress on the boundary in the vicinity of the contact, which physically means that the solid will generate a reaction to this external force. This will affect the geometrical positions of the boundary point itself and of the adjacent internal point (because they are used in the calculation of the stress on the boundary): the boundary point and its adjacent internal neighbour will need to be positioned in such a way as to maintain the reactive stress distribution on the boundary, required to balance off the applied external boundary force.

Meanwhile, the dynamics of the adjacent internal point is governed by the dynamic equation, of which one of the terms is $\sigma^{ij}\big|_j$. As we have discussed earlier, the calculation of this derivative will involve the use of the two sets of neighbouring points for each direction, one of which will be our boundary point. Because of this and the fact that the two points (one internal and the boundary one) are now repositioned to maintain the required stress on the boundary, the term $\sigma^{ij}\big|_j$ will "feel" this disturbance and thus distribute it further inside the body, in the same direction as the external force.

To summarise, after we have had a disturbance on the boundary due to the external boundary force, the boundary condition will ensure this disturbance is felt by the internal layer of points, "adjacent" to the boundary. Once it has reached the internal points, it will be further propagated by the dynamic equations throughout the solid.

It should be noted that it is not quite precise to say that the term $\sigma^{ij}\big|_j$ "feels" the disturbance and then distributes it. The term $\sigma^{ij}\big|_j$ is itself a vehicle of the distribution; it responds to a disturbance. Its task is specifically to distribute any imbalance between individual stresses σ^{ij} in each direction, regardless of what has caused this imbalance.

Theoretically, this system of equations is all we need to be able to solve our modelling problem. It completely identifies how to find the positions of all points of the solid after the start of the experiment, subject to the given internal (volume) and external (boundary) forces, by providing a set of equations all the points (internal and boundary) must satisfy at all times.

Practically, it is very difficult to solve this system of equations in the given form. The first equation, the dynamic second order differential equation, is normally solved numerically by stepping through time. This process, however, is impossible to apply for the boundary points. The boundary condition is static - it is a static coupled system of equations, binding all the boundary points, as well as the adjacent internal points. These adjacent internal points are involved simultaneously in the dynamic equation and in the static boundary condition.

It is clear, that to be able to solve these equations practically, we will need to find a way to solve the outlined difficulties.

We will show later, in section 6.2 where we will discuss the numerical implementation, that it is indeed possible to create a numerical solution scheme, which will use the boundary condition but will only be dealing with dynamic differential equations.

2.9 Limitations of the linear Classical Theory of Elasticity

The classical theory of elasticity provides a rigorous and comprehensive theoretical framework for solving modelling problems involving elastic solids.

However, problems still exist. To get a better understanding of the nature of the issues still requiring attention, it is useful to recall what sort of object we are to model. In our computer modelling we will deal with highly elastic bodies, in fact most of the items in the computer-simulated world will be elastic (or totally rigid). The objects will interact, move around, rotate and be expected to deform significantly away from their original shapes during impacts and/or under the action of applied external forces.

Indeed, we can formulate precisely two of the most evident requirements for a successful modelling theory:

- The theory must be able to deal with the objects moving and rotating freely. There should be no restrictions on how far a body can move away from its original position or how much it can rotate.

- The theory must provide the means to model large scale object deformations. Some limitations can be placed on the range of the deformations, but the number of these limitations, as well as the nature of the limitations, must be minimal.

As we will show in section 4.2.1, the classical theory of elasticity in its linear form is unable to meet these two basic requirements. It establishes a very useful and rigorous framework, but it requires further development in order to become useful for practical modelling problems. Section 4.2 provides a comprehensive and detailed discussion of the limitations of the linear classical theory of elasticity. The following two sections provide a simple overview of these limitations.

2.9.1 Globally small displacement gradients

In the linear classical theory of elasticity the notion of the displacement field is introduced. It is defined as a displacement vector field, describing the positions of the points of the body as it deforms away from its original shape.

The strain tensor is introduced first in its original precise form (through the metric tensors of the body in its deformed and undeformed shape). Later its expression through the displacement field is derived:

$$\varepsilon_{ij} = \frac{1}{2}\left(U_i\big|_j + U_j\big|_i + U^m\big|_i U_m\big|_j\right) \tag{2.265}$$

This is still a precise formula for the strain tensor. Later, however, an assumption of globally small displacement gradients is made and the non-linear term is dropped. The following *linearized* strain is used instead:

$$\varepsilon_{ij} = \frac{1}{2}\left(U_i\big|_j + U_j\big|_i\right) \tag{2.266}$$

Dropping the non-linear term and using the linearized strain seriously restricts the range of the deformations, which the theory can handle adequately: the displacement gradient field must be globally small, or, in other words, the body must not deform or move away too much from its original shape.

It could be argued that using the non-linear strain (where the non-linear term is preserved) will solve this problem and lift the restriction on the displacements. Practically, however, it will lead to a very poorly conditioned system that is difficult to solve. Additionally, the required computational load on the numerical solver will dramatically increase as well: it will need to additionally compute the product of various covariant derivatives of the displacement field (when calculating the non-linear term) and then numerically calculate its covariant derivative to obtain the derivative of stress.

We will show in the description of our theory of FlexyMatter that it is possible to remove the restriction of the globally small displacement gradient and at the same time still deal with the linear strain thus avoiding the use of this non-linear term and all the problems that it brings with it. It is possible to stay within the *same* framework as developed in the linear classical theory of elasticity but at the same time dramatically increase the applicability of the theory. Our approach is based on a modification of the notion of the displacement field.

Practically, it is easy to see just how restrictive the requirement of the globally small displacement gradient is. Let us for example consider an elastic rod fixed at one end. We apply some force at the opposite end and want to find how the rod will deform. From practical experience we know that the deformation of the loose end of the rod will depend on the applied force and the length of the rod. Let us have a fixed applied force (for example gravity). We can see the deformation will depend on the length of the rod only: the longer the rod the more it will bend. The restriction of the globally small displacement gradients in this example will mean that there is a limit on the length of the rod: the theory will only be able to model adequately very short rods. As the length of the rod increases (and therefore the deformation), the theory will quickly lose its precision (as well as possibly stability) and will lead to physically unrealistic results.

2.9.2 Rigid body motion and rotation

We say a body moves rigidly if it does not deform during its motion, and quasi-rigidly if it deforms only slightly during its motion.

We will show in section 4.2 that the linear classical theory is not suitable for modelling rigid or quasi-rigid bodies, where the body is expected to move around or rotate too much. In both cases it will lead to dramatic global increase in the displacement field gradient, which is unacceptable.

This is a very serious limitation of the applicability of the theory. For example, a body that moves in a straight line and slowly rotates, without any external influences or forces and no actual deformation cannot be modelled using the linear classical theory at all.

Several modifications were proposed to lift this particular limitation. Most of them are considered in detail in the chapter 3. Unlike the other methods, our approach removes this limitation while still staying within the framework of the linear classical theory.

2.10 Summary

In this chapter, we presented the classical theory of elasticity: a comprehensive theory for modelling elastic solids.

We started with the precise formulation of the modelling problem from the mathematical point of view. We then quickly derived a basic law of motion for a solid by just using Newton's second law of motion. From this derivation, it became clear that there exists an internal force (called the elastic force) that acts on each internal point and which is a volume force. We have obtained a number of useful properties of this elastic force by applying simple reasoning.

In order to obtain a complete expression for the internal elastic force, a number of additional mathematical concepts were introduced: curvilinear coordinate systems, coordinate transformations, covariant and contravariant base vectors, tensors, covariant derivatives and so on.

Then the notion of the strain tensor, as a geometrical measure of the body's deformation, was introduced and its expression through the displacement field was obtained. In order to proceed, we made an assumption of a globally small displacement gradient field and introduced linearized strain. The notion of the stress tensor, as a measure of the internal distribution of forces, was then presented.

The constitutive equations, binding the stress and the strain were introduced. These equations encapsulate the internal material properties of the solid. Practical expressions of the elastic moduli E^{ijlm} were presented for isotropic materials.

By this stage we had enough information to introduce the final equation for the internal elastic force. Combining the basic law of motion we derived earlier with the expression for

the elastic force, we obtained a complete set of dynamic equations for the internal points of the solid.

The boundary condition, the equations relating the internal stress and the external boundary force, was introduced to provide an equation for the boundary points of the solid. Combining this boundary condition for the boundary points with the dynamic equations of motion for the internal points, allowed us to obtain a complete system of equations for the evolution of the solid under the action of the applied internal and external forces.

We have identified that the theory of elasticity provides a rigorous and comprehensive framework for modelling elastic deformations. However, in its linear form the theory of elasticity suffers from several limitations that make it difficult to apply the theory in practice. Most important is the assumption of the globally small displacements gradients - this dramatically restricts the range of applicability of the theory. This assumption excludes not only large elastic deformations, but also rigid body displacements and rotations.

Our approach not only addresses and removes all the outlined limitations of the linear classical theory of elasticity, but also makes several important simplifications, particularly important when solving the problem numerically. At the same time, it still uses the classical framework and is therefore mathematically rigorous. It can be treated as a natural extension of the linear classical theory of elasticity that extends its applicability into the domain of applied computer modelling.

3. Established Practical Modelling Methods

As we have already mentioned in section 1.2 there is large amount of practical research accumulated over the last twenty years in the area of applied modelling. The problem of finding a fast, flexible and comprehensive modelling technique to apply to a variety of practical situations has captured the attention of the research community.

Many practical modelling methods and techniques have been proposed, ranging from very specific methods (such as modelling cloth [5,6,15,22,38,39,43,46,53,58,63,64]) to generic methods ([7,12,89]). In this chapter we take a close look at the range of practical methods developed over the last twenty years. We concentrate on the techniques that present practical interest from the point of view of this thesis, that is, well-established practical techniques for modelling elastic solids.

We begin with a simple classification. We divide the whole range of applied modelling methods into two groups: local modelling methods and global modelling methods. For each of these groups we provide a discussion of the group, together with a detailed description of the typical practical methods that fall into the group.

3.1 Local Modelling Methods

Local modelling methods are those where the system of equations describing the evolution of all the points of a solid is formulated purely in local terms, that is, the terms that operate over infinitesimal neighbourhoods of the points of the solid. In such methods, the information used to determine the evolution of any given point is only drawn from the immediate neighbourhood of the point in question. Because the evolution is described

purely in local terms, the information about any disturbance in the solid is spread in waves travelling with a finite speed away from the source of the disturbance. For example, if a solid is pushed, the points in the immediate vicinity of the contact feel the disturbance first (via the increased strain). This creates an area of increased strain, which spreads away from the zone of contact in the opposite direction with a finite speed. As a result of this, a point far from the area of contact will not feel the disturbance immediately, but only after a certain time, after the information about the disturbance reaches its infinitesimal neighbourhood.

For internal points of the solid, the systems of equations for local modelling methods have the following generic form:

$$\ddot{\mathbf{R}}\rho = \mathbf{F} + \mathbf{E} \tag{3.1}$$

That is, the acceleration $\ddot{\mathbf{R}}$ of the internal point multiplied by the density of the material at the point ρ is equal to the sum of the resultant of the external forces $\mathbf{F}$ and the elastic force $\mathbf{E}$. $\mathbf{E}$ is the resultant of the surface forces transmitted via the surface of the infinitesimal neighbourhood of the point from the rest of the solid, or in other words it is the resultant of the action of the rest of the solid on the point through its infinitesimal neighbourhood.

Various practical methods are distinguished by the way they calculate the elastic force $\mathbf{E}$, since the acceleration $\ddot{\mathbf{R}}$, the density ρ and the resultant of the external forces $\mathbf{F}$ are external factors and have the same clear meaning in any practical method.

Fundamentally, there are two different types of methods of calculating the elastic force $\mathbf{E}$; all the practical modelling methods are based on one of these. The first one is based on the notion of scalar energy and the second one is based on the notion of the displacement field. These two types are discussed in detail below.

3.1.1 Energy based methods

In the energy based methods a notion of the energy of deformation E is introduced. This energy E is a non-negative scalar function over all the points of the solid, which has the following properties:

- E increases when the solid is deformed and can therefore be regarded as a (scalar) measure of deformation.

- E is a minimum when the solid is in equilibrium under the action of the applied forces.

It can then be shown that the elastic force is expressed through the variational derivative of this scalar energy E:

$$\mathbf{E} = -\frac{\delta E}{\delta \mathbf{R}} \tag{3.2}$$

A number of practical methods are energy based: Terzopoulos et al [7, 12], Gudukbay et al [36], Eberhardt et al [58], Provot [53] and Magnenat-Thalmann et al [15,22,43,46,63,64]. The energy based approach is one of the most popular approaches in applied modelling.

We will now consider several typical examples in detail.

The primary formulation by Terzopoulos et al [7, 12]

The primary formulation developed by Terzopoulos et al [7, 12] can be considered the first rigorous treatment of the applied modelling problem. It has served as a basis for many other researchers (including, for example, Gudukbay et al [36] and Magnenat-Thalmann et al [15,22,43]).

In this formulation, in the case of a three-dimensional solid, the potential energy E of the solid's deformation is given as follows:

$$E = \iint \left\| \mathbf{G} - \mathbf{G}_0 \right\|^2 dx^1 dx^2 dx^3 \tag{3.3}$$

Here, $\mathbf{G}$ is a matrix, consisting of the metric tensor elements $G_{ij} = \frac{\partial \mathbf{R}}{\partial x^i} \cdot \frac{\partial \mathbf{R}}{\partial x^j}$, describing the metric properties of the deforming solid. $\mathbf{G}_0$ is a matrix with the metric tensor elements, describing the metric properties of the undeformed solid; $\|\circ\|$ is a weighted matrix norm.

The elastic force at a point is then approximated with the following expression, which is an approximation of the variational derivative of the energy E:

$$\mathbf{E} = -\frac{\partial}{\partial x^i}\left(\eta_{ij} \left(G_{ij} - G_{ij}^0 \right) \frac{\partial \mathbf{R}}{\partial x^j} \right) \tag{3.4}$$

Here η_{ij} determine the material properties of the solid at each point: $\eta_{ii}, i = 1,2,3$ determine the resistance to stretching in three different directions and $\eta_{ij}, i \neq j$ determine resistance to shear deformation in various directions. These parameters are adjusted on a trial and error basis; there is no explicit connection to any other parameters commonly used in engineering (such as the modulus of elasticity, Poisson ration or modulus of rigidity).

The primary formulation by Terzopoulos et al is an advanced energy based modelling technique. The authors provide energy expressions not only for the three-dimensional case, but also separately for one and two-dimensional cases. They also used a simple energy dissipation term to account for the loss of energy during the solid's evolution, which should also act as a useful stabilising factor in a numerical implementation.

However problems still exist. In addition to a number of parameters that need adjusting in the model itself on a trial and error basis, the method offers no clear description on how to deal with surface forces. Since in any practical simulation, surface friction and interaction of solids are the issues that invariably need to be addressed, this lack of clear description of how to deal with surface forces makes the primary formulation practically incomplete in this respect.

Additionally, the dissipation term suggested by Terzopoulos et al is too simplistic a method to account for the energy loss. In the form suggested by the authors the energy dissipation term leads to a physical system where any moving object loses energy at all times. This will lead, for example, to an undeformed solid given an initial velocity without any external forces to eventually stop instead of continuing to move at a constant velocity. In other words, the simulation system would have a sort of universal "ether friction", which lacks a physical equivalent in the real world. A more sophisticated means to deal with the internal energy dissipation is required (see section 6.3 for one possible treatment of the internal energy dissipation we used in our theory of FlexyMatter).

A particle system model by Eberhardt et al [58]

Eberhardt et al [58] proposed a non-continuous approach – a particle based model for modelling the appearance of cloth (represented geometrically as a two dimensional surface). They suggested treating the cloth as a collection of interacting weighed particles.

In this formulation the coordinates identifying the positions of the particles together with their velocities are treated as degrees of freedom of the system. Therefore for a system of n particles, there are $6n$ degrees of freedom: $R_i^1, R_i^2, R_i^3, \dot{R}_i^1, \dot{R}_i^2, \dot{R}_i^3, i = \overline{1,n}$

The potential V is then introduced:

$$V = \sum_{i=\overline{1,n}} \left(E_{pot}^i + E_t^i + E_s^i + E_b^i \right) \tag{3.5}$$

Here E_{pot}^i is the potential energy of point number i (due to its spatial position relative to Earth), E_t^i is the tension energy (which develops when the material is stretched/squashed), E_s^i is the shearing energy (which develops when the material suffers lateral distortion) and E_s^i is the bending energy (which develops when the material is bent).

The elastic energy is then obtained from the potential V:

$$E^i = -\frac{\partial V}{\partial R^i} \tag{3.6}$$

The following expressions for the individual energies were used:

$$E_{pot}^k = m_k g R_k^3$$

$$E_t^k = \begin{cases} \sum_{i=1}^{4} \frac{1}{2} C_{i,t_1} \left(|\mathbf{p}_0 - \mathbf{p}_i| - d_i - h_{t_i} \right)^3, |\mathbf{p}_0 - \mathbf{p}_i| \geq d_i \\ \sum_{i=1}^{4} \frac{1}{2} C_{i,t_2} \left(|\mathbf{p}_0 - \mathbf{p}_i| - d_i - h_{t_i} \right)^5, |\mathbf{p}_0 - \mathbf{p}_i| < d_i \end{cases} \tag{3.7}$$

$$E_s^k = \sum_{i=1}^{4} \frac{1}{2} C_s \left(\varphi_i - \frac{\pi}{2} - h_{s_i} \right)^2$$

$$E_b^k = \sum_{i=1}^{2} \frac{1}{2} C_b \left(\psi_i - \pi - h_{b_i} \right)^2$$

Here $\mathbf{p}_0$, $\mathbf{p}_i$ and angles φ_i and ψ_i are determined as shown in Figure 3.1 below:

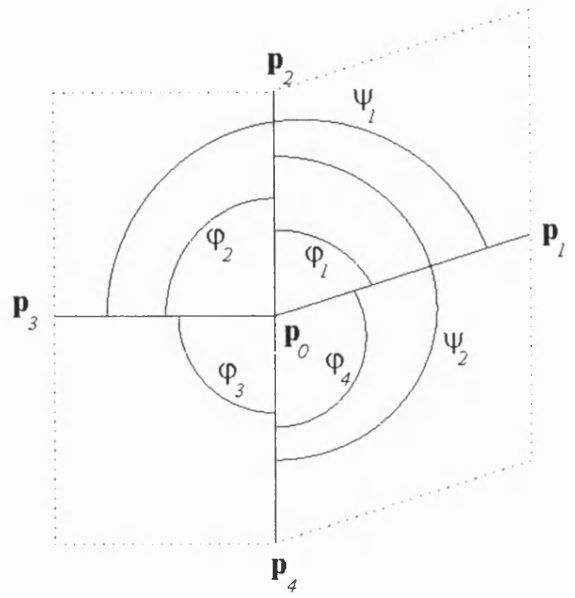


Figure 3.1: Illustration for the calculation of the potential energy at a point

Constants d_i are the distances between the corresponding points in the undeformed body and the constants C, h are determined using the Kawabata plots.

The expression for the elastic force can be obtained explicitly by differentiating the appropriate expressions of the energies. This can be done manually or by symbolic differentiation using computer algebra systems.

The particle system developed by Eberhardt et al is a simple and efficient model for modelling two-dimensional objects (such as cloth). It is easy to understand and it affords a straightforward numerical implementation.

However, there are several practical problems with this system, especially when one attempts to use it to model three-dimensional objects. The system has been formulated specifically for the two-dimensional case – its expansion to the three-dimensional case is possible but not straightforward. The energy expression in this case will become significantly more complex, especially due to the usage of angles in the energy expressions. The procedure of adjusting the myriad of parameters in the model is based on the two-dimensional Kawabata tests; thus in the three-dimensional case, another procedure will have to be devised, which is most likely to be a trial and error approach. Finally, the model lacks any means to account for internal energy dissipation, which may present some practical stability problems in numerical implementations.

A mass-spring model by Provot [53]

Provot [53] developed a physically based technique for modelling cloth. He treated the cloth as a collection of massive particles, connected by springs to their neighbours in various ways (see Figure 3.2).

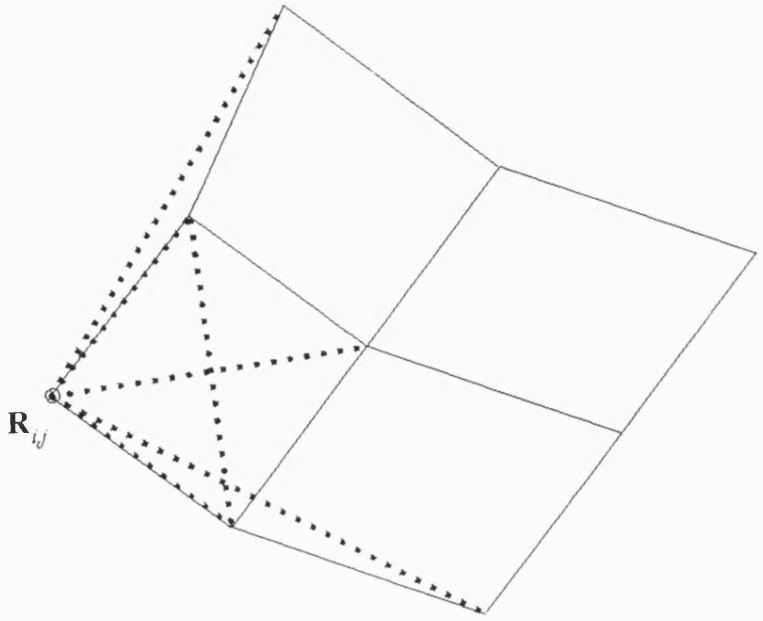


Figure 3.2: A node with the set of virtual springs connecting it to its neighbours

There are three types of springs:

- Structural springs – the two springs that link $\mathbf{R}_{i,j}$ with $\mathbf{R}_{i+1,j}$ and $\mathbf{R}_{i,j+1}$.
- Shear springs – the two springs that link $\mathbf{R}_{i,j}$ with $\mathbf{R}_{i+1,j+1}$ and $\mathbf{R}_{i+1,j}$ with $\mathbf{R}_{i,j+1}$.
- Flexion springs – the two remaining springs that link $\mathbf{R}_{i,j}$ with $\mathbf{R}_{i+2,j}$ and $\mathbf{R}_{i,j+2}$.

The elastic force was then approximated explicitly using Hooke's law for ideal springs:

$$\mathbf{E}(i,j) = -\sum_{k,l} K_{i,j,k,l}\left[\mathbf{l}_{i,j,k,l} - l^0_{i,j,k,l}\frac{\mathbf{l}_{i,j,k,l}}{\left\|\mathbf{l}_{i,j,k,l}\right\|}\right] \tag{3.8}$$

Where $K_{i,j,k,l}$ is the stiffness of the spring linking the current node $\mathbf{R}_{i,j}$ with its neighbour $\mathbf{R}_{k,l}$, $\mathbf{l}_{i,j,k,l}$ is the vector connecting the current node with the corresponding neighbour, $l^0_{i,j,k,l}$ is the distance between the current node and its corresponding neighbour in the undeformed configuration.

This particular formulation may look quite different from the ones developed by Terzopoulos et al and Eberhardt et al, but on careful examination it may be noted that in reality they are very similar. The energy expression for the node $\mathbf{R}_{i,j}$ can be formulated as follows:

$$E(i,j)=\sum_{k,l}\frac{1}{2}K_{i,j,k,l}\left[\left\|\mathbf{l}_{i,j,k,l}\right\|-l^{0}_{i,j,k,l}\right]^{2} \tag{3.9}$$

Where $\|\circ\|$ is the usual Euclidean vector length (the square root of the sum of squares of the coordinate differences). If we then write this expression in the coordinate form and take the partial derivative by $R^{1}_{i,j}$, $R^{2}_{i,j}$ and $R^{3}_{i,j}$, we will obtain the force expression (3.8) used by Provot.

This mass-spring model by Provot is a simple technique for modelling cloth and cloth-like objects, which is easy to understand and implement. However, this mass-spring model is very specific to modelling two-dimensional objects because of the way it has been constructed (due to the use of "flexion springs"). Additionally, it suffers from all the problems outlined for the previous particle system by Eberhardt et al [58], including the following:

- Any attempt to expand the model to deal with three-dimensional objects would lead to a complex expression for the elastic force with a large number of parameters that need adjusting on a trial and error basis.
- It lacks any means to account for internal energy loss.

We will now proceed to consider another group of local modelling methods, those based on the notion of the displacement field.

3.1.2 Methods based on the notion of the displacement field

We have shown that all the local modelling methods are based on the generic equation of evolution (3.1) describing the evolution of the points of a solid with respect to time. We have shown that different individual methods in the group of local modelling methods provide different ways to evaluate the elastic force term $\mathbf{E}$ – the force that arises inside the solid in reaction to its deformation.

In the previous section we have discussed a group of methods that are energy based. For all these methods the elastic force is evaluated as a covariant derivative of the scalar energy E as shown in (3.2).

A distinctly different approach to the evaluation of the elastic force is based on the notion of the displacement field, which we will discuss in this section. In this approach, we introduce the notion of a reference frame, which is the solid in its undeformed configuration somehow positioned in space, and as a consequence, the notion of a displacement field – for each point of the deforming solid its displacement is the difference between its position vector in the deforming solid and the corresponding position vector in the reference frame (Figure 3.3).

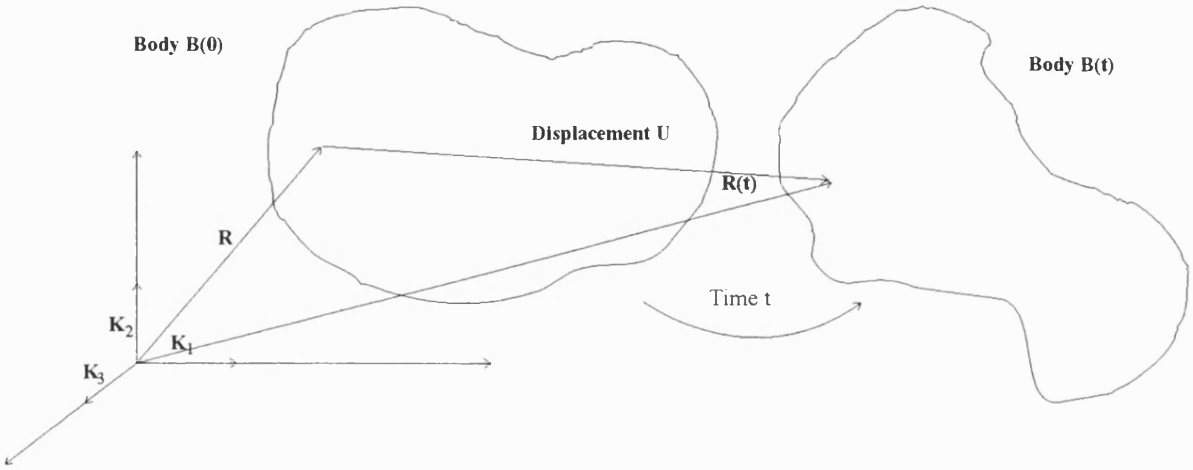


Figure 3.3: Definition of the displacement field

Here, the reference frame is $\mathbf{B}(0)$ and the deforming body is $\mathbf{B}(t)$ (this notation is chosen because normally the reference frame is the solid itself at the start of the experiment).

Having introduced a reference frame and therefore a displacement field, a coordinate system on the reference frame is introduced with the basis vectors given by the following expressions:

$$\mathbf{G}_i = \frac{\partial \mathbf{R}}{\partial x^i} \tag{3.10}$$

We can now represent the displacement field in this coordinate system:

$$\mathbf{U} = U^i \mathbf{G}_i \tag{3.11}$$

The elastic force $\mathbf{E}$ is expressed via the displacement field and is normally represented in the coordinate system $\mathbf{G}_i$.

This approach is the basis of the classical theory of elasticity and other derived methods such as our theory of FlexyMatter and the theory of Hyper-Matter by Smith, Paddon [89,90] (when $\delta \to 0$).

The classical theory of elasticity was established over a hundred years ago and provides a fundamental framework for modelling elastic deformations. Due to its long established history, it has been applied extensively in engineering to various practical problems, including propagation of sound and other elastic waves, structural design to assess the strength of constructions and so on.

However, the theory in its classical form is not well suited for modelling highly elastic solids (see section 4.2 for a detailed discussion of the reasons for this). To become a successful practical theory for modelling highly deformable elastic solids, it needs to be modified. Two such modifications have been proposed to date: the theory of Hyper-Matter by Smith, Paddon in 1992 [89,90] and our theory of FlexyMatter. Both theories suggest improvements to the classical theory itself in order to make it suitable for practical modelling tasks.

We will now proceed to consider the group of global modelling methods.

3.2 Global Modelling Methods

In the global modelling methods the systems of equations describing the evolution of the individual points of a solid contain both global quantities, such as the solid's centre of mass or its inertia tensor, and purely local quantities, such as the density of the material at the point.

In the local modelling methods, the information about any disturbance in the solid travels around in waves; it is "passed" from one infinitesimal neighbourhood to its adjacent one through their joint boundary. This is no longer the case for global modelling methods, because the global quantities are present in the description of the evolution of local points of a solid. For example, a disturbance that leads to a change in the geometrical shape of the solid will have an immediate affect on all of the solid's global quantities, such as the centre of mass and the tensor of inertia. Therefore, the information about such a disturbance can spread instantly throughout the solid.

There are two practical examples of the global modelling methods: the theory of Hyper-Matter by Smith, Paddon [89,90] (when $\delta = 1$) and the hybrid formulation by Terzopoulos et al [12], which we will consider in detail below.

The hybrid formulation by Terzopoulos et al [12]

After developing the primary formulation [7] as a method for modelling (highly) elastic bodies, Terzopoulos et al extended their approach with an additional technique, which they called the hybrid formulation. The motivation was to develop a method for modelling semi-rigid bodies, that is, bodies that deform only slightly and therefore behave almost like rigid bodies. The primary formulation was most suitable for modelling highly elastic bodies and was not adequate when the body was semi-rigid. This is due to numerical difficulties involved in enforcing the "rigidness" in the primary formulation; the more rigid the solid is the faster the energy has to grow when it is deformed away from its equilibrium state. This leads to strong internal forces even for small deformations and, as a result, serious practical difficulties in the numerical scheme.

In their hybrid formulation Terzopoulos et al introduce a reference frame – the body in its undeformed shape. The reference frame is allowed to move, rotate and displace freely according to the applied forces. A point on the deformed body is then presented as a sum of its position vector on the reference frame plus the displacement vector (similar to the classical theory of elasticity). See Figure 3.4 below:

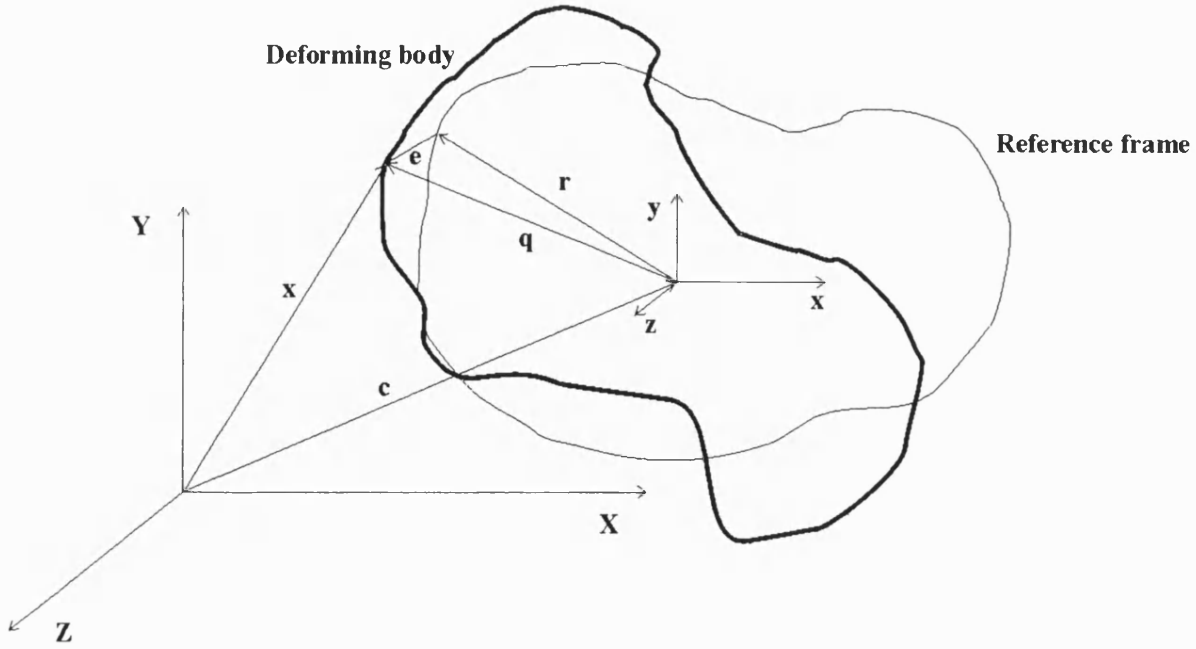


Figure 3.4: Deforming body and its reference frame

In the hybrid formulation the equations of motion have the following form:

$$\begin{cases} \frac{d}{dt}(m\mathbf{v}) + \frac{d}{dt}\int_\Omega \rho\dot{\mathbf{e}}\,du + \int_\Omega \gamma\dot{\mathbf{x}}\,du = \int_\Omega \mathbf{F}\,du \\ \frac{d}{dt}(\mathbf{I}\boldsymbol{\omega}) + \frac{d}{dt}\int_\Omega \rho\mathbf{q}\times\dot{\mathbf{e}}\,du + \int_\Omega \gamma\mathbf{q}\times\dot{\mathbf{x}}\,du = \int_\Omega \mathbf{q}\times\mathbf{F}\,du \\ \frac{d}{dt}(\rho\dot{\mathbf{e}}) = \mathbf{F} + \mathbf{E} - [\rho\dot{\mathbf{v}} + \rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{q}) + 2\rho\boldsymbol{\omega}\times\dot{\mathbf{e}} + \rho\dot{\boldsymbol{\omega}}\times\mathbf{q}] - \gamma\dot{\mathbf{x}} \end{cases} \tag{3.12}$$

Where:

$$\begin{aligned} \mathbf{v} &= \frac{d}{dt}\int_\Omega \rho\mathbf{x}\,du \\ \boldsymbol{\omega} &= \dot{\boldsymbol{\theta}} \\ \dot{\mathbf{x}} &= \mathbf{v} + \boldsymbol{\omega}\times\mathbf{q} + \dot{\mathbf{e}} \end{aligned} \tag{3.13}$$

Here $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ is the orientation of the coordinate system attached to the reference frame relative to the global coordinate system; γ is the damping coefficient; $\mathbf{I}$ is the inertia tensor; m is the total mass of the solid ($m = \int_\Omega \rho\,du$) and the elastic force $\mathbf{E}$ is calculated via the variational derivative of a scalar energy, as in the primary formulation.

As can be seen in (3.12), we no longer have a simple system with the generic form of (3.1). Indeed, because of the fact that we now have a reference frame, which is allowed to translate and rotate under the action of the applied forces, we have:

- Two additional equations (the first and the second in (3.12)) to govern the evolution of the reference frame. The first equation describes the evolution of its centre of mass and the second – the evolution of its rotational orientation.
- Two more terms in the last equation in (3.12): $-[\rho\dot{\mathbf{v}} + \rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{q}) + 2\rho\boldsymbol{\omega}\times\dot{\mathbf{e}} + \rho\dot{\boldsymbol{\omega}}\times\mathbf{q}]$, which consists of the sum of pseudo forces due to the fact that the coordinate system attached to the reference frame is not inertial; and $-\gamma\dot{\mathbf{x}}$, which is simply a damping term. This damping term is used extensively by Terzopoulos et al [7, 12] as a simple measure to account for the energy dissipation during solid's evolution.

Comparing this to the local modelling methods, one can note greatly increased complexity of the system. We now have three entities to deal with: the centre of mass of the reference

frame, its orientation and the relative deformation. Additionally, since the equations for the points of the solid are formulated in terms of the displacement (denoted by the vector $\mathbf{e}$) away from the reference frame, there are many additional force terms (pseudo-forces) that have to be added to account for the fact that the reference frame is allowed to move and rotate freely. The above issues, plus the lack of an adequate method for dealing with the internal energy dissipation and surface forces (as in the case of the primary formulation), make the hybrid formulation problematic in actual use. In fact, it means that the hybrid formulation is difficult to implement as a practical modelling technique.

It can even be said that the hybrid formulation is actually a step in the wrong direction, especially when we note that all the added complexity is there simply to handle rigid body translations and rotations, which can be solved by other, more efficient theories (such as our theory of FlexyMatter and the theory of Hyper-Matter when $\delta = 1$).

The theory of Hyper-Matter by Smith, Paddon [89,90] (when $\delta = 1$)

In 1992 Smith, Paddon [89,90] developed a theory of Hyper-Matter, based on the classical theory of elasticity. But instead of having a static reference frame, as in the classical theory, they suggested using a different reference frame for different points in the solid. Smith, Paddon therefore introduced a notion of a Hyper-Matter – a three-dimensional object defined over a six dimensional domain. Hyper-Matter contains in itself the deforming body as well as all of its reference frames simultaneously at all times.

Instead of allowing the reference frame to move and rotate freely and thus having to explicitly deal with its equations of motions as a rigid body (as in the hybrid formulation proposed by Terzopoulos et al [7,12]), Smith, Paddon suggested simply realigning the reference frame with the deforming body at each time step, using a simple alignment procedure.

In the theory of Hyper-Matter Smith, Paddon introduced a scalar parameter $0 < \delta \leq 1$, which he called the domain of linearity. When $\delta = 1$ there is only one single reference frame at any time step, which is the undeformed solid aligned translationally and rotationally with the deforming solid. Once it is aligned, the displacement field can be calculated in the same way as in the classical theory of elasticity. All the related quantities (strain, stress and the elastic force) are then calculated in the same way as in the classical theory.

The basic equation of motion for the solid is therefore the same as in the classical theory of elasticity (3.1), where the elastic force is calculated via the covariant derivative of stress.

We can see that this formulation offers clear advantages over the hybrid formulation by Terzopoulos et al [12]. It is much simpler and easier to understand and implement. As opposed the hybrid formulation by Terzopoulos et al, the theory of Hyper-Matter is based on the extension of the classical theory of elasticity and offers a simple procedure to realign the reference frame at every time step. This automatically allows for the deforming solid to move and rotate freely without any added complexity to the basic equations of motion (3.1).

However, in general the theory of Hyper-Matter as presented by Smith, Paddon suffers from several shortcomings. In particular, Hyper-Matter is an entity presented as a three-dimensional solid defined over six dimensions, incorporating in it the deforming solid and all of its reference frames at all times. This required formal expansion of all basic mathematical entities (such as tensors) and operations involving them (such as taking a derivative) to be formally defined over a six-dimensional domain. This expansion into six dimensions is unnecessarily complex, as strictly speaking it is only an infinitesimal six dimensional neighbourhood for every point of the solid that is ever practically used (see section 5. where the parameters of the calculation of the displacement field are discussed for a detailed discussion of this topic).

3.3 Summary

In this chapter, we have looked at a range of practical modelling methods developed over the last twenty years. We have classified them into two broad groups: local modelling methods and global ones, judging by the presence of the global terms (such as the centre of mass and the inertia tensor) in the description of the evolution of the points of a solid.

The global modelling methods are best suited for modelling semi-rigid solids - solids that behave almost like rigid bodies. The group consists of only two practical methods: the theory of Hyper-Matter by Smith, Paddon [89,90] and the hybrid formulation by Terzopoulos et al [12]. The hybrid formulation is very complex, while the theory of Hyper-Matter (when $\delta = 1$) offers a simpler and easier to implement method for handling semi-rigid bodies.

On the other hand, most methods from the local group of modelling methods (all but the classical theory of elasticity) are best suited for modelling highly elastic solids. There is

normally a clear distinction between the methods that can handle highly elastic solids and those that deal with the semi-rigid bodies. Normally, in the methods that are best suited for handling highly elastic solids, the closer the solid is to a perfectly rigid body, the more difficult it is to solve the system of equations of motion; special numerical techniques must often be used. This is due to numerical problems (so called stiffness problem).

The group of local modelling methods offers a wide choice of practical methods. However, problems still exist:

- Those that are based directly on the classical theory without serious modifications cannot model large deformations because of problems in the classical theory itself.
- The theory of Hyper-Matter by Smith, Paddon is a comprehensive modelling technique but it has an unnecessarily complex formulation, written in an extremely terse and compact fashion, which does not help its acceptance in the wide research community.

A large number of practical modelling methods from the group of local modelling methods are energy based and as a result:

- Most suffer from the problem of the existence of large numbers of parameters that need adjusting on a trial and error basis, with no clear link to the parameters widely used in engineering (such as modulus of elasticity, Poisson ratio).
- All but the hybrid formulation by Terzopoulos et al are primarily formulated for two-dimensional objects, such as cloth; energy expressions for the three-dimensional case become very complex indeed.
- None offers a clear description on how to handle surface forces and the internal energy dissipation. This is a key problem because such practical issues as friction and contact between solids are all related to a distribution of surface forces over the solid's boundary. The lack of means to deal with the internal energy loss can be a practical problem in a numerical implementation.

Our theory of FlexyMatter has been developed in response to these problems and offers a comprehensive practical modelling solution, capable of handling highly deformable solids. It is relatively simple, rigorous (derived from the classical theory of elasticity) and well adapted for practical implementations (offers practical techniques for handling surface

forces, rigorous treatment of internal energy dissipation and so on). It can model a wide range of deformations: from very floppy to quasi-rigid within the same framework and without the need for the special treatment of the case of quasi-rigid solids.

4. The need for a new modelling theory

4.1 Introduction

In chapter 2 we presented the classical theory of elasticity – a classical framework for modelling the behaviour of elastic solids. We identified that the vector function $\mathbf{R}(x^i,t)$ representing the deformed body will satisfy the following equation:

$$\ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \mathbf{E}(x^i,t), \forall x^i \in \operatorname{int} I, \forall t > 0 \qquad (4.1)$$

Here $\rho(x^i,t)$ is the material density, $\mathbf{F}(x^i,t)$ is the external volume force and $\mathbf{E}(x^i,t)$ is the internal (elastic) force, which arises due to the deformation of the body. The equation itself is given in the global Cartesian coordinate system K.

For all internal points, the elastic force is given by a simple expression:

$$\mathbf{E}(x^k,t) = \sigma^{ij}(x^k,t)\big|_j \mathbf{G}_i(x^k) \qquad (4.2)$$

For boundary points, the situation is different: there is no dynamic equation, but rather a static one. The boundary condition is given as follows:

$$\mathbf{S}(x^k,t) = \sigma^{ij}(x^k,t) n_j(x^k,t) \mathbf{G}_i(x^k) \qquad (4.3)$$

Here $n_j(x^k,t)$ is a unit normal vector to the boundary surface at the point x^k, pointing outwards.

The stress tensor σ^{ij} is expressed ultimately through the covariant derivatives of the displacement field U_i, defined as the displacement of every point away from its original position:

$$\sigma^{ij} = E^{ijkl}\varepsilon_{kl}$$

$$\varepsilon_{kl} = \frac{1}{2}\left(U_k|_l + U_l|_k\right) \tag{4.4}$$

Thus defined strain is an approximate linearized strain, which differs from the accurate one by the lack of a non-linear term (see (2.224)).

4.1.1 The classical theory of elasticity as a fundamental framework for modelling deformable solids

The classical theory of elasticity was established more than a hundred years ago as a fundamental framework for modelling elastic solids. It was shown that, in a simple case of homogenous isotropic solids, the equations describing elastic deformations could be expressed via the derivatives of the displacement field with only two constants, representing intrinsic properties of the elastic material. The most commonly used constants in applications are the modulus of elasticity E (representing the resistance of the material to stretching) and the Poisson ratio ν (representing the degree of lateral bulging/thinning when squashed/stretched). Many practical experiments were conducted to calculate the values of these constants for real materials [1].

The classical theory of elasticity has been successfully applied in such specific areas as the propagation of sound and other elastic waves, structural design to assess the strength of constructions and so on. It has been shown to be a simple and relatively easy to use modelling theory.

In the 1980's, with the great rise of available computational power, the problem of modelling *highly* elastic materials, in particular cloth, received increased attention from the research community. This sort of modelling problem is normally difficult to solve explicitly; an approximate numerical solution scheme must be used, and more powerful computers offered the required computational resources for the task.

The classical theory of elasticity in its linear form seemed unlikely to cope with modelling highly elastic bodies. It was hard to imagine how the model, that proposes to use the body at the beginning of the experiment as a reference frame and expresses the internal elastic force via the displacements away from it using only the linear displacement gradients, can effectively cope with modelling solids moving freely, rotating and deforming at the same time. The solid to be modelled can displace and rotate arbitrarily, perhaps even without

actually deforming. The selection of the body in its original position as the reference frame also seems rather artificial, and there is no clear understanding of the effect that a solid's rotation and translation may have on the internal elastic force, which explicitly depends on the displacement field.

These apparent difficulties lead to a quiet dismissal view of the linear classical theory of elasticity as a modelling theory capable of providing an adequate description of the deformation of highly deformable solids, which are allowed to move and rotate without any restriction. It was merely assumed to be only applicable to problems where small actual deformations are assumed beforehand and no arbitrary rotation and translation is allowed.

Indeed, in the first attempts to model highly deformable bodies, such as cloth, energy methods were used (Feynman [6], Weil [5], Terzopoulos [7, 12]). In these methods scalar energy of deformation was introduced, with its variational derivative being the elastic force. The energy methods became very popular, due to their relative simplicity. However, in our opinion, using the energy methods constituted an unnecessary departure away from the classical theory of elasticity. In addition, the energy methods themselves had problems. In the calculation of the energy a large number of parameters were used; it was clear that these parameters somehow related to the material properties of the material, but how to set them for real materials was not clear at all. Also, the energy methods were only applied to modelling cloth. It was only suggested how the same methodology could also be used to model three-dimensional solids (Terzopoulos [7, 12]). The implementation, however, would have been a very difficult task: the expression for the energy of a three-dimensional solid would have been very complex indeed, with even more parameters that would need to be adjusted somehow. It was also unclear how to deal with surface forces (like traction, wind and air resistance) and internal energy dissipation.

Some researchers made attempts to apply the classical theory of elasticity (for example, Aono [18]). Very specialised techniques were proposed, suitable for a limited range of applications (for example, wrinkle propagation in cloth in Aono's case). No modifications to the linear classical theory to make it more suitable for practical applications were proposed.

It is interesting to note that no attempts were made to find out why the linear classical theory of elasticity cannot be effectively applied in practical modelling problems; no one made attempts to analyse the problem in detail. We believe strongly that if this analysis had been

made, the modifications to the classical theory proposed in our theory of FlexyMatter would have been discovered sooner.

We believe that the linear classical theory of elasticity has not lost its relevance since its inception and provides an excellent foundation for a modern and rigorous theory of modelling highly deformable solids, such as the theory of FlexyMatter presented in this chapter. There is no reason in our view to ignore many years of experience and the wealth of research accumulated simply because the classical theory cannot be applied directly to modelling highly deformable solids. In this chapter we will show that it is possible to make several simple modifications to the linear classical theory so that the extended theory (the theory of FlexyMatter) will provide a simple, mathematically rigorous and flexible applied theory for modelling highly deformable solids.

In the next section we will present a detailed analysis of the problems arising from the direct application of the classical theory of elasticity to modelling moving and rotating deformable solids. We will then proceed to outline one possible solution to these problems: the global alignment approach (section 4.3.1). We will show that this approach, popular among some researchers, offers only a partial solution to the problems with the classical theory and introduces new unwanted side effects. The theory of FlexyMatter will then be presented as a better alternative to the global alignment approach. It will be shown to resolve successfully the limitations of the classical theory of elasticity and it will be presented formally in the section 5.. Finally, we will present the practical solution of the modelling problem, including numerical solution of the differential equations, followed by the examples.

4.2 Analysis of the limitations of the classical theory of elasticity

The main problem in the theory lies in the expression for the internal elastic force $\mathbf{E}$ (and all other quantities on which it depends: stress, strain and the displacement field) – the response of the solid to the deformation. This elastic force takes part in the following set of equations:

$$\begin{cases} \ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \mathbf{E}(x^i,t), \forall x^i \in \operatorname{int} I \\ \mathbf{S}(x^k,t) = \mathbf{E}_s(x^i,t), \forall x^i \in \partial I \\ \mathbf{E}(x^i,t) = \sigma^{ij}\big|_j(x^k,t)\mathbf{G}_i(x^k) \\ \mathbf{E}_s(x^i,t) = \sigma^{ij}(x^k,t)n_j(x^k,t)\mathbf{G}_i(x^k) \end{cases} \tag{4.5}$$

All the other elements of the equations are known: $\ddot{\mathbf{R}}(x^i,t)$ is the acceleration of the point of the solid, $\rho(x^i,t)$ is its density at the point, $\mathbf{F}(x^i,t)$ and $\mathbf{S}(x^k,t)$ are the external forces, $n_j(x^k,t)$ is the outward normal to the boundary, $\mathbf{G}_i(x^k)$ are the covariant base vectors from the reference frame, and $\sigma^{ij}(x^k,t)$ is the stress tensor.

The elastic force $\mathbf{E}$ is a function of the displacement field $\mathbf{U}(x^k,t)$, calculated from the stationary reference frame to the point in its current position $\mathbf{R}(x^i,t)$:

$$\mathbf{U}(x^k,t)=\mathbf{R}(x^i,t)-\mathbf{R}(x^i) \tag{4.6}$$

The reference frame stays fixed throughout our experiment - we choose it to be coincident with the solid at the start of the experiment.

Let us temporarily forget how the elastic force is calculated and just look at the dynamic equation and the boundary condition:

$$\begin{aligned} \rho\ddot{\mathbf{R}} &= \mathbf{F}+\mathbf{E} \\ \mathbf{S} &= \mathbf{E}_s \end{aligned} \tag{4.7}$$

We have the inertial force $\rho\ddot{\mathbf{R}}$, which is due to the volume element having certain mass, equal to the external volume force $\mathbf{F}$ plus the elastic (reaction) force $\mathbf{E}$. We know that by definition $\mathbf{E}$ is the force that arises as a result of a solid's deformation; it "attempts" to restore the balance by moving the point to a position where the overall deformation is reduced. (From this we can make a simple conclusion that the elastic force $\mathbf{E}$ must be zero if the body is not deformed.)

The boundary condition is static; it describes the fact that the internal surface force $\mathbf{E}_s$ transmitted through the boundary and the external surface force $\mathbf{S}$ applied to the boundary must be in balance at all times.

Clearly, the dynamic equation and the boundary condition are generic; they are not the result of the theory of elasticity, but simply relate to Newton's second law. The classical theory of elasticity provides a practical procedure for the evaluation of the internal boundary force $\mathbf{E}_s$ and the internal elastic volume force $\mathbf{E}$ for real solids. There are other theories

that provide a practical way to evaluate the elastic force. For example, in the energy methods, $\mathbf{E}$ is calculated as a variational derivative of the scalar internal energy.

In other words, (4.7) expresses universal laws, while the classical theory of elasticity describes one method of how $\mathbf{E}$ and $\mathbf{E}_s$ can be practically evaluated.

As with any other method, it has limitations - in certain cases, the elastic forces approximated according to the classical theory of elasticity are calculated with an error. Depending on applications, this error can be acceptable but it can also become critical. In animation, for example, the precision of the model is not so important as long as the results look plausible.

The aim of this chapter can now be formulated as follows: to investigate errors that the classical theory introduces to the elastic volume force $\mathbf{E}$ and the elastic boundary force $\mathbf{E}_s$.

4.2.1 Errors introduced by the linear classical theory of elasticity in the evaluation of $\mathbf{E}$ and $\mathbf{E}_s$

First, note that when the body deforms only slightly away from the reference frame (its original position), there is a large amount of practical experience that illustrates that the classical theory of elasticity provides an adequate evaluation for $\mathbf{E}$ and $\mathbf{E}_s$. This is true regardless of the orientation of the deforming body and the reference frame in space (as long as their relative positions are preserved), since the elastic force is evaluated via the displacement field.

We now proceed to consider a solid that can move, rotate and deform arbitrarily, and thus can no longer be considered to be close to its position at the beginning of the experiment (regardless of whether it is actually deformed or not, that is, it can translate and/or rotate and/or deform).

Assume that we are at some time into our experiment; our body is arbitrarily displaced, rotated and possibly deformed relative to the reference frame (or, for the translation and rotation, we can equally say that the reference frame is rotated and/or translated relative to the deforming body). In other words, we need to deal with three types of independent transformations: rigid body translation, rigid body rotation, and actual deformation. As usual, by rigid body transformation we mean a transformation, which preserves distances between any two points (does not distort its topology).

At this stage we are not concerned with the question of how a general transformation can be separated into these three components; we will consider this in the next section, where we discuss our proposed modifications. In fact in this section we will only consider relative translations and rotations. Since the displacement field is calculated away from the reference frame and is therefore sensitive to the relative position of the reference frame, our aim here will be to examine what effect relative translation and rotation of the deforming body (or, inversely, the reference frame) will have on the calculation of the elastic forces.

Neither we are concerned with how or why the deforming body gets translated and/or rotated relative to its reference frame. We are thinking abstractly: we have a deforming body and its reference frame somewhere in space. We can calculate the displacement field and therefore the elastic forces for the current positions of the body and its reference frame. What we want to discover is: if we now simply displace the body for some fixed vector or rotate it around a fixed axis, what effect this will have on the elastic forces $\mathbf{E}$ and $\mathbf{E}_s$. We must ask the question – will they be affected or not? Practical experience and common sense suggests that $\mathbf{E}$ and $\mathbf{E}_s$ will not be affected. If we have a body, its reference frame and thus calculate the distribution of internal (elastic) forces in a body, it is obvious that if we simply displace or rotate the reference frame, it must not affect the distribution of the elastic forces inside the body (because the body itself does not change in any way). Formally speaking, both $\mathbf{E}$ and $\mathbf{E}_s$ must be invariant to translation and rotation of the deforming body (or, equivalently, the reference frame) because they measure pure deformation.

We will now show that, in fact, both $\mathbf{E}$ nor $\mathbf{E}_s$ are invariant to translational transformation, but neither $\mathbf{E}$ nor $\mathbf{E}_s$ are invariant to rotational transformation. As we have noted above, we can equivalently consider translation and/or rotation of the deforming body relative to its reference frame or translation and/or rotation of the reference frame relative to the deforming body.

We will subject the solid to simple transformations and consider the effect this would have on the displacement field (because the elastic forces $\mathbf{E}$ nor $\mathbf{E}_s$ depend linearly on the first and second derivatives of the displacement field).

4.2.2 Pure translation

We assume that the whole transformation consists only of a pure rigid body translation. Since the body is not deformed, $\mathbf{E}$ in this case must be zero everywhere in the body. The displacement field is simply a constant vector:

$$\mathbf{U}(x^k,t)\equiv\mathbf{C} \tag{4.8}$$

Since $\mathbf{E}$ and $\mathbf{E}_s$ are expressed via the derivatives of the displacement field, we can see clearly that they will be identically zero:

$$\mathbf{U}_{,j}(x^k,t)\equiv\mathbf{0}\Leftrightarrow U^i\big|_j(x^k,t)\equiv 0 \tag{4.9}$$

Since both elastic forces $\mathbf{E}$ and $\mathbf{E}_s$ are expressed via the derivatives of the displacement field, from this we can conclude that they will also be identically zero everywhere in the body. Pure translation therefore does not have any effect on the elastic forces $\mathbf{E}$ and $\mathbf{E}_s$.

4.2.3 Pure rotation

We will now show that both $\mathbf{E}$ and $\mathbf{E}_s$ are *not* invariant to the pure rigid body rotation of the reference frame. We will show that as the relative position of the reference frame changes (rotating around a fixed axis), the body starts to develop a non-zero $\mathbf{E}_s$ and $\mathbf{E}$.

Let us suppose our full transformation consists of a pure rotation around a fixed axis. As there is no deformation, there must be no internal forces, thus both $\mathbf{E}$ and $\mathbf{E}_s$ must be identically zero.

We will represent the rotational transformation by a matrix $\mathbf{A}(\tau)$ (τ here is merely a parameter) and assume that the centre of the rotation is $\mathbf{R}_0$.

The position vector of the body can be expressed via its reference frame:

$$\mathbf{R}(x^k,t)=\mathbf{A}(\tau)(\mathbf{R}(x^k)-\mathbf{R}_0)+\mathbf{R}_0 \tag{4.10}$$

And the displacement field is therefore:

$$\mathbf{U}(x^k,t)=\mathbf{A}(\tau)(\mathbf{R}(x^k)-\mathbf{R}_0)-(\mathbf{R}(x^k)-\mathbf{R}_0)=(\mathbf{A}(\tau)-\mathbf{I})(\mathbf{R}(x^k)-\mathbf{R}_0) \tag{4.11}$$

Here, $\mathbf{I}$ is the unit matrix.

As we know, the internal elastic volume force $\mathbf{E}$ is expressed through derivatives of the stress tensor, and the internal elastic boundary force $\mathbf{E}_s$ is expressed through the stress tensor itself. The stress tensor in turn is expressed via the first derivatives of the displacement field.

To understand the effect this pure rotation will have on $\mathbf{E}$ and $\mathbf{E}_s$, we need to calculate the first and the second derivatives of the displacement field. The first derivative is:

$$\mathbf{U}_{,j}(x^k,t)=(\mathbf{A}(\tau)-\mathbf{I})(\mathbf{R}_{,j}(x^k))=(\mathbf{A}(\tau)-\mathbf{I})\,\mathbf{G}_j(x^k) \tag{4.12}$$

And the second derivative is:

$$\mathbf{U}_{,jl}(x^k,t)=(\mathbf{A}(\tau)-\mathbf{I})\,\mathbf{G}_{j,l}(x^k) \tag{4.13}$$

As we have mentioned, the first derivative relates to the elastic boundary force $\mathbf{E}_s$ and the second derivative of the displacement field relates to the elastic volume force $\mathbf{E}$.

From the expressions for the first and second derivatives we can clearly see that both $\mathbf{E}$ and $\mathbf{E}_s$ depend directly on the rotational matrix $\mathbf{A}(\tau)$ and in general are not zero. Moreover, the degree of the "distortion" grows with the rotation, reaching a maximum at a certain value of the parameter τ.

Let us examine now in which cases $\mathbf{E}$ and $\mathbf{E}_s$ are identically zero.

For $\mathbf{E}$:

$$\begin{gathered}\mathbf{U}_{,jl}(x^k,t)=0\Leftrightarrow\\ (\mathbf{A}(\tau)-\mathbf{I})\,\mathbf{G}_{j,l}(x^k)=0\Leftrightarrow\\ \mathbf{A}(\tau)=\mathbf{I}\vee\mathbf{G}_{j,l}(x^k)=0\end{gathered} \tag{4.14}$$

If $\mathbf{A}(\tau)=\mathbf{I}$, there is no rotation. From the second condition, it follows that the base vectors are constant:

$$\mathbf{G}_{j,l}(x^k)=0\Rightarrow\mathbf{G}_j(x^k)=\mathbf{G}_j \tag{4.15}$$

This is only true for Cartesian or Affine coordinate systems.

For $\mathbf{E}_s$:

$$\mathbf{U}_{,j}\left(x^k,t\right)=0 \Leftrightarrow$$

$$\left(\mathbf{A}(\tau)-\mathbf{I}\right)\mathbf{G}_j\left(x^k\right)=0 \Leftrightarrow \tag{4.16}$$

$$\mathbf{A}(\tau)=\mathbf{I}$$

This means that the only way the boundary elastic force can be zero is if there is no rotation of our body.

Combining the conditions for both $\mathbf{E}$ and $\mathbf{E}_s$, we can see that unless there is no rotation, the elastic forces will be calculated with errors; the more relative rotation of the body and its reference frame is accumulated throughout the evolution, the larger will be the error in the evaluation of the elastic forces, with the effect particularly profound in the case of the boundary elastic force $\mathbf{E}_s$. Moreover, we can also see that unless we are working with a Cartesian or Affine coordinate system on the reference frame, the internal elastic force $\mathbf{E}$ will also be calculated with an error.

We can now see why for small deformations the elastic forces $\mathbf{E}$ and $\mathbf{E}_s$ are evaluated adequately. In the case of uniformly small deformations we would have a uniformly small deformation field, and the relative rotation of the deforming body will be insignificant (that is, $\mathbf{A}(\tau)$ will be very close to the unity tensor $\mathbf{I}$).

As a conclusion, we can state that in the case of a general deformation (which can be separated into pure translation, pure rotation and pure deformation), the pure rotation of the reference frame will have a negative effect on the calculation of the elastic forces $\mathbf{E}$ and $\mathbf{E}_s$, particularly on the boundary elastic force $\mathbf{E}_s$. We have shown that the errors introduced depend directly on the magnitude of the rotation.

Having outlined the precise nature of the errors the elastic forces will accumulate if the reference frame is rotated relatively to the deforming body, we can now proceed to outline how our proposed extensions to the classical theory of elasticity will eliminate these errors. We will discard the whole notion of the static reference frame; each node will have its own reference frame relative to which the forces will be evaluated.

4.3 Possible modifications to the classical theory of elasticity

We have shown that the procedure of calculating the elastic forces $\mathbf{E}$ and $\mathbf{E}_s$, prescribed by the classical theory of elasticity, is invariant to pure rigid body translations. We have also shown that it is not the case with rigid body rotations: both $\mathbf{E}$ and $\mathbf{E}_s$ will be calculated with errors if the reference frame and the deforming body are rotated relative to each other.

It is clear that it is this sensitivity to the relative rotation of the deforming body and its reference frame that is responsible for limiting the applicability of the classical theory of elasticity. It is also clear that any successful extension of the classical theory of elasticity must aim to resolve this issue, and the success of the extension will be judged mostly on how well it has addressed and resolved this problem concerning the sensitivity to the relative rotation.

We will show that, having eliminated this problem of errors in $\mathbf{E}$ and $\mathbf{E}_s$, we can obtain a theory that is capable of modelling both freely moving and rotating solids and highly deformable solids.

We will proceed in two stages:

(1) In this chapter, we will present one obvious but not entirely adequate path to address this sensitivity to relative rotation, by allowing the reference frame to move rigidly by continually globally aligning it with the deforming body. This will expand the theory to deal with rigid body translations and, more importantly, rotations.

This is an obvious extension for the classical theory of elasticity; it has been used by many researchers (most notably Terzopoulos [7]). But, in our opinion, it is inadequate. It is the wrong way forward; it leads to unnecessary complication of the theory, does not resolve all the issues entirely and, more importantly, introduces its own problems. We will discuss these in more detail in the next section.

(2) As the second stage, in the next chapter, we will show a new way to expand the classical theory. We will consider local rotational aligning at every point, instead of the global alignment. We will show that this way we both allow the body to move and rotate freely, and allow it to deform almost arbitrarily, whilst still remaining within the framework of the classical theory of elasticity.

4.3.1 One possible extension: global alignment of solid and its reference frame

The most straightforward approach to expand the classical theory of elasticity to resolve some of its most obvious shortcomings is to have a moving reference frame instead of the stationary one. As the body moves, rotates and deforms, we can account for its global translation and rotation by continuously aligning the reference frame and the deforming body.

Indeed, we can split the body's overall transformation into translation, rotation and actual deformation in the following way:

- First, we calculate the average position of all the points of the reference frame, which is its geometrical centre, and that of the body in its current position. We will obtain two vectors: their difference is the rigid body translation (that is, the average translation).

- Secondly, we calculate the average rotation of all the points of the reference frame, relative to a globally fixed axis, and that of the body in its current position. We will obtain two vectors: their difference is the rotation (that is, the average relative rotation).

- Thirdly, to obtain the actual deformation, we translate the reference frame using the translational transformation and rotate it using the rotation transformation. The difference between the body in its current position and thus the transformed reference frame is the actual deformation.

Each of these transformations is independent: we can apply them in any sequence we wish, and still arrive at the body in the same displaced position.

It is natural to call the actual deformation obtained this way the "pure" deformation. However, as we will show, it is generally incorrect and misleading. We will show that the displacement field obtained from the aligned reference frame will be "better" than that obtained from the stationary reference frame because it will not contain rigid body translations and rotations. However, due to the existence of the local rotations, it is still wrong to call the deformation obtained in this way a "pure" deformation.

Figure 4.1 illustrates an example of a moving and rotating solid. In examples, we will use simple one-dimensional[1] solids in this chapter for simplicity.

In this example, we have a solid that is displaced through a distance of $\mathbf{D}(t)$ and rotated through an angle $\Phi(t)$. For simplicity, we consider a solid that does not deform.

As we have described, the rigid body translation $\mathbf{D}(t)$ is obtained by calculating the difference between the position vectors $\mathbf{C}(t)$ and $\mathbf{C}$, which are the geometric centres of the deforming solid and the reference frame respectively.

After the solid and its reference frame are aligned translationally, the rotational alignment angle is obtained by calculating the average rotation of the deforming body and the average rotation of the reference frame. We can use these to obtain the relative rotation of the deforming solid and the reference frame.

We associate a matrix $\mathbf{A}$ with this coordinate transformation.

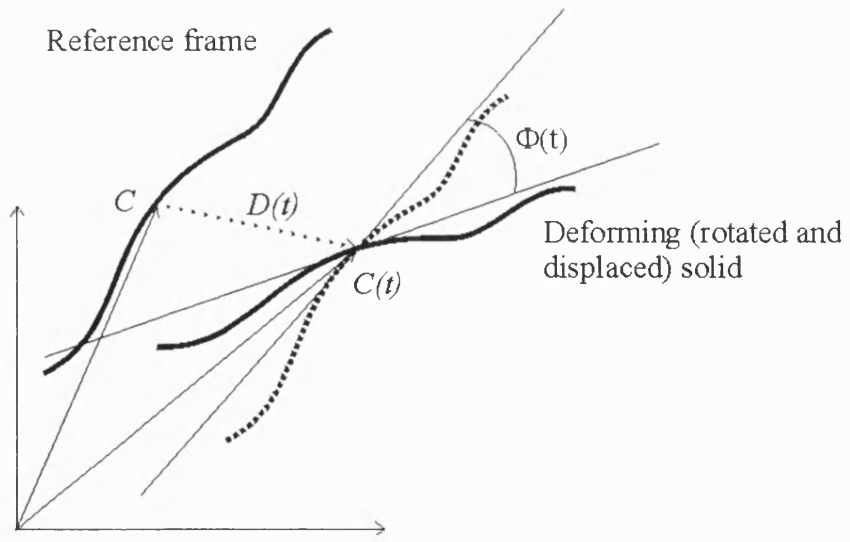


Figure 4.1: Relative rotation of the deforming solid

Here, $\mathbf{C}(t)$ and $\mathbf{C}$ are the geometrical centres of the deforming solid and the reference frame respectively, $\Phi(t)$ is the angle of the relative rotation and $\mathbf{D}(t)$ is the rigid body displacement.

---

[1] A one-dimensional solid is physically a three-dimensional solid, whose linear dimensions in two directions are negligibly small. Therefore, the dimension of the problem does not change in these examples: it is still a three-dimensional problem.

The rotation matrix $\mathbf{A}$ corresponds to the transformation, rotating every vector through an angle of $\Phi(t)$:

$$\mathbf{A} = \mathbf{A}(\Phi(t)) = \mathbf{A}(t) \tag{4.17}$$

As we know, the only way to eliminate errors in the calculation of the elastic forces (both boundary and volume) due to the existence of the relative rotation (displacement $\mathbf{D}(t)$ will not cause any errors) is to have $\mathbf{A}(t) = \mathbf{I}$, which means that we must eliminate relative rotation.

In the example, shown Figure 4.1, to achieve this we need to rotate the reference frame through angle $\Phi(t)$ using the transformation $\mathbf{A}(t)$. That is, instead of considering the original reference frame, we must consider the rotated copy of it. Even though we know that the rigid body displacement $\mathbf{D}(t)$ does not cause any errors, we can also align the bodies translationally by moving the reference frame through distance $\mathbf{D}(t)$. It is this rotated and displaced copy of the reference frame that we will use to calculate the displacement field.

It is clear from this simple example that we must discard the notion of a stationary reference frame. There is no particular reason why the reference frame must be the same at all times, we can clearly see that it can and indeed should be different at each instance of time of our experiment, aligned rotationally with the deforming body. The alignment is determined by global properties of the deforming body and the reference frame (the average rotation and translation).

This kind of global alignment resolves a number of limitations of the classical theory: the body can now translate and rotate without any limitations. Since the body and its reference frame are continuously aligned, the displacement field will not be distorted by these translations and rotations.

It is not, however, a complete solution to all the existing problems and, in our opinion, is generally an inappropriate way to address these problems. This is because there are two serious issues with the global aligning approach:

- First of all, the deformation part of the transformation as separated out in this scheme is not pure deformation. Thus, there is still a significant limit on the degree of the allowable deformation with a similar problem of the elastic forces being

distorted by small local rotations. We will demonstrate it in detail with an example in the subsection *Problem 1: deformation is not pure* below.

- It introduces several global quantities such as centre of mass and average rotation. This leads to very peculiar and unwanted side effects. We will consider these in subsection *Problem 2: new problems* below.

One way to address the above issues is to modify the global alignment approach. However, we maintain that it is inappropriate to modify the global alignment approach for this purpose because it would lead to substantial complication of the theory without clear improvements. Instead, it is more appropriate to get back to the classical theory of elasticity and start anew. Our new approach, as we will show, not only removes the limitations on the rigid body translations and rotations (as the global alignment method does) but also gives a correct separation of the pure deformation, significantly increasing the degree of the allowed deformations of the solid.

Problem 1: deformation is not pure

We will now consider the first issue with the global alignment approach in more detail. Take the following example, illustrated below in Figure 4.2. Our body deforms in such a way that it stays aligned to the reference frame at all times. In this case, we will not need to do any additional alignment; we can calculate the forces immediately.

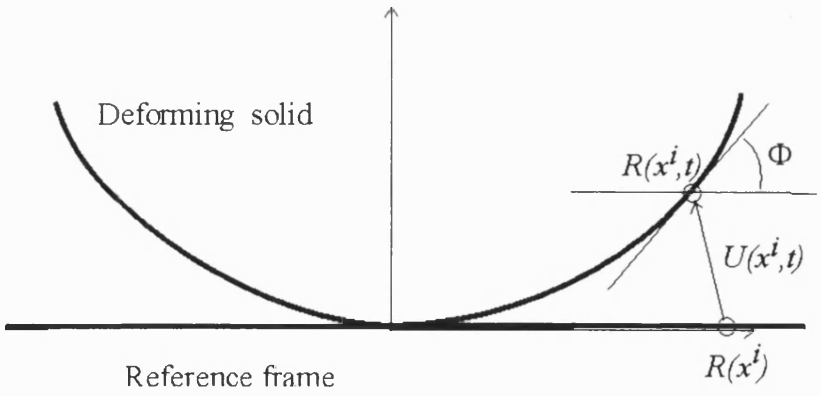


Figure 4.2: Deformation with alignment

We can clearly see that the deforming body and the reference frame are globally aligned rotationally. This should mean that the displacement field calculated from the stationary reference frame would represent the pure deformation as defined by the global alignment approach. We will now show that this is not true.

We consider a point x^k that is far from the centre of alignment. Figure 4.3 below shows a magnified version of the right part of Figure 4.2.

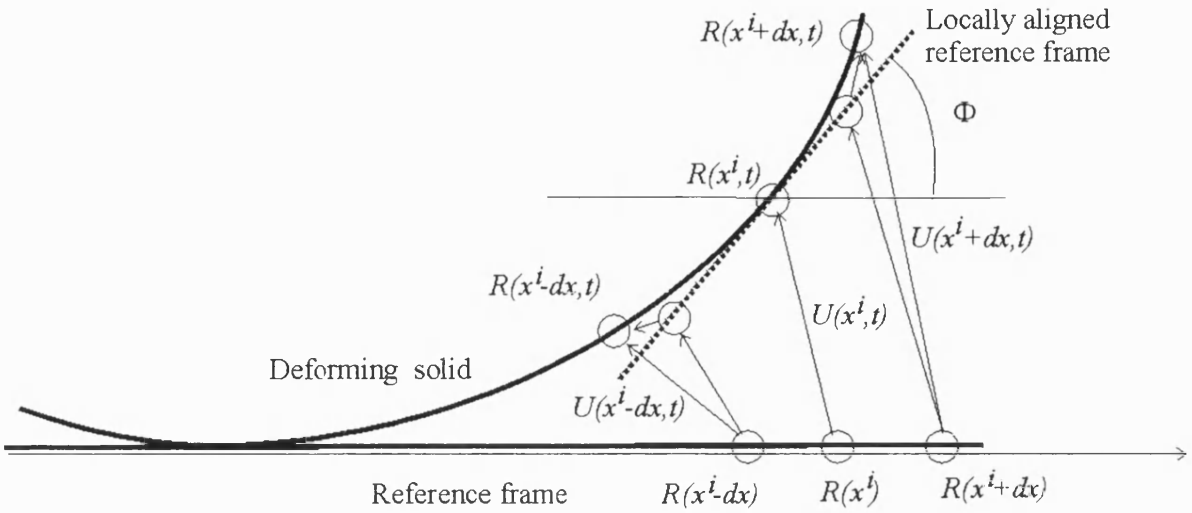


Figure 4.3: Locally misaligned deforming body and reference frame

Remember that the elastic forces are functions of the first and second derivatives of the displacement field. The derivatives of the displacement field are by definition local quantities; they measure the rate of change of the quantity in the immediate neighbourhood of the point.

From Figure 4.3 we can see that the displacement field $\mathbf{U}(x^k,t)$ does not represent pure deformation: instead it has a component that is due to the rotation of the local neighbourhood (it also has a translational component but because this has no effect on the derivative we may ignore it). We can see that $\mathbf{U}(x^k,t)$ can be represented as a sum:

$$\mathbf{U}(x^k,t)=\mathbf{U}_{rot}(x^k,t)+\mathbf{U}_{def}(x^k,t) \tag{4.18}$$

Here, $\mathbf{U}_{rot}(x^k,t)$ is the component of the displacement due to the rotation of the local neighbourhood and $\mathbf{U}_{def}(x^k,t)$ is the remainder (which does, in fact, represent pure deformation).

When we form derivatives of the displacement field, we can see that will still have an error component due to the local rotation:

$$\mathbf{U}_{,j}(x^k,t)=\mathbf{U}_{rot,j}(x^k,t)+\mathbf{U}_{def,j}(x^k,t) \tag{4.19}$$

$$\mathbf{U}_{,jl}(x^k,t) = \mathbf{U}_{rot,jl}(x^k,t) + \mathbf{U}_{def,jl}(x^k,t)$$

We have shown earlier that neither $\mathbf{U}_{rot,j}(x^k,t)$ nor $\mathbf{U}_{rot,jl}(x^k,t)$ will be zero unless the reference frame and deforming solid are aligned, in this case locally, that is, that the angle $\Phi = 0$.

Even though the deforming body and its reference frame are constantly aligned globally, as suggested by the global alignment approach, the local elastic forces will still be calculated with errors due to the rotations of the local neighbourhoods of the points. This demonstrates that the deformation as separated by the global alignment method is not pure; it contains small aberrations due to these local rotations.

In fact, this example shows that it is generally impossible to achieve a global separation into translation, rotation and deformation (with one single displacement vector and rotational matrix), where the deformation can be considered "pure". Whatever the global alignment, local neighbourhoods of some points will still be misaligned and therefore the elastic forces will be calculated with errors.

We can see the kind of limitations on the degree of the global deformation that this approach can handle.

Problem 2: new problems

The global alignment approach introduces several unwanted side effects. One of the most important of them is the effect of the global quantities on the evaluation of elastic force. This problem arises because the global alignment approach mixes together global quantities, such as average rotation and average position, with inherently local quantities, such as strain, stress and elastic force.

Let us consider the average rotation of the deforming solid. This is clearly a function of all points of the solid. The reference frame is aligned according to the average rotation of the deforming body, thus the alignment of the reference frame is a function of all points of the deforming body. The displacement field is calculated away from the aligned reference frame. Therefore, we may conclude that the displacement field at any point as well as the elastic forces (volume and boundary) are functions of all the points of the deforming body simultaneously.

This leads to the following peculiar fact: a change in the position of a point far away from a given point can have a profound effect on the elastic force calculated at the given point. This is clearly an artefact introduced by the global alignment approach, which has no real physical explanation and is therefore an artefact introduced by the approach.

4.4 Summary

In this chapter we presented the arguments for a new theory for modelling elastic solids, based on the classical theory of elasticity.

We began the chapter with an introductory note on the status of the classical theory of elasticity as a fundamental framework for modelling deformable solids. In this section we also point out that the classical theory itself is not suitable for modelling highly elastic solids moving and rotating freely. Thus most of the practical techniques for modelling highly elastic bodies developed over the last twenty years are based on other approaches, most notably the energy methods. These other approaches have been used without any attempts to investigate the shortcomings of the classical theory in detail and to examine the possibility of extending it to deal with a full range of applied modelling problems, including highly elastic solids moving and rotating freely.

We therefore dedicated section 4.2 to a thorough investigation of the limitations of the linear classical theory of elasticity. We first noted that it is the elastic forces (the internal force $\mathbf{E}$ and the boundary force $\mathbf{E}_s$) in the equations of motion that are erroneously evaluated in certain circumstances. Since the elastic forces depend ultimately on the derivatives of the displacement field, we conclude that we need to look carefully at what happens to the first and second derivatives of the displacement field when we subject the solid to certain simple transformations (such as pure rigid body translation and rotation). As the result of applying these transformations and noting the effect they have on the derivatives of the displacement field, we notice that it is rotation that is the root of the problems, or, more precisely, the local misalignment of the point's infinitesimal neighbourhood and that of the corresponding point from the reference frame.

In section 4.3 we discussed one possible modification to the classical theory of elasticity: the global alignment approach, where the solid and its reference frame are constantly globally aligned. This method improves the classical theory from the point of view of applied modelling. But, as we point out at the end of section 4.3.1, the global alignment approach only addresses some of the limitations of the linear classical theory (namely its

inability to model solids moving and rotating freely), while introducing some of the new artefacts connected with the usage of global quantities during the alignment. As such, we conclude, the global alignment approach is not the best way forward.

In the next chapter, we will present a different approach to remove the limitations of the linear classical theory of elasticity: the local alignment approach, used in our theory of FlexyMatter. In this approach a differently aligned copy of the reference frame is used to evaluate the elastic forces at each point at all times. This local alignment approach solves all the problems outlined in section 4.2 and, as we will show, is a better way to extend the linear classical theory of elasticity to be applicable in the modern applied modelling tasks.

5. The Theory of FlexyMatter

5.1 Introduction

In this chapter we will provide a detailed introduction into our theory of FlexyMatter – a novel approach to solving the major limitations of the linear classical theory of elasticity, presented in 4.2.1.

We will begin by introducing the notion of local alignment of the reference frame and the deforming body. In this way we will remove all the associated problems with the erroneous evaluation of the elastic forces, and, as a consequence of this, we will remove all the limitations on the allowed rigid body motion and rotation of the deforming solid.

As we know, the derivatives of the displacement field and therefore the elastic forces are purely local quantities. They only take into account the geometric positions of the neighbouring points. This makes perfect physical sense: the elastic force is a response to the local geometrical deformation of the solid away from its original (local) shape, and thus should only depend on the points in the immediate vicinity of a point. Any approach that makes the elastic force dependent on other points far away from the point at which the force is evaluated is fundamentally flawed and should be rejected.

As we showed earlier, when the deforming body and the reference frame are misaligned, the error introduced into the calculation of the elastic forces depends directly on the rotation matrix, such that the error is zero when the rotation matrix is unity. We have also seen that there is no dependency on the centre of rotation; it is the rotation itself that causes the errors. The elastic force is local in nature and the error depends only on the relative misalignment of the solid and its reference frame. Thus, only the local misalignment of the solid and the reference frame can and will disturb the elastic force.

The elastic force at a point is not affected by distant points, regardless of whether or not the solid and the reference frame are aligned at these distant points. It is the alignment of immediate neighbourhood of a point that is of paramount importance for precise evaluation of the elastic force at that point.

The global alignment approach described in 4.3.1, globally aligns the solid and its reference frame. Being globally aligned the solid can globally rotate and translate freely, without any restrictions; the deformations, however, are still restricted to be globally small. In the theory of FlexyMatter, alignments are made locally, and thus different for each point of the solid. See Figure 5.1 for an illustration of this process.

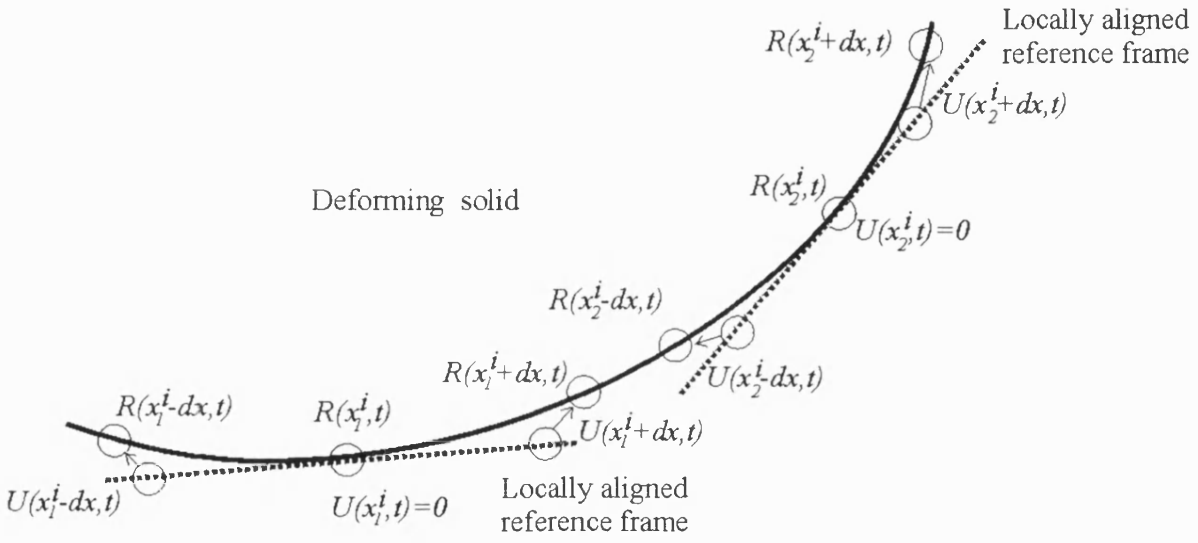


Figure 5.1: Local alignments

The local alignment places no restrictions on the rotations and translations of the local infinitesimal neighbourhoods of the points. We can conclude that by applying the same reasoning as in the global alignment method: translations have no affect on the force evaluation and the rotations are accounted for by continuous alignment of the local neighbourhood.

If we consider the neighbourhoods of all points, their union will be the solid itself. Therefore, if any of the local neighbourhoods can rotate and translate freely, the *whole* of the body will also be able to rotate and translate freely. Therefore, the local alignment approach is at least as general as that using the global alignments.

Moreover, because we only work with a local neighbourhood when evaluating the elastic forces for the point, there is no restriction on the degree of global deformation the whole body can undergo. Whatever happens outside the local infinitesimal neighbourhood of a

point has no effect on any of the quantities that are evaluated at that point: strain, stress, and volume and boundary elastic forces.

More formally, the displacement field in this case will be a function of both the point at which the alignment has been made and the point at which it is actually taken:

$$\mathbf{U} = \mathbf{U}_{x_0^i}\left(x^k, t\right) \tag{5.1}$$

Here, x_0^i is the point at which the solid and reference frame are aligned, $x^k \in \Omega\left(x_0^i\right)$ is the point at which it is evaluated and $\Omega\left(x_0^i\right)$ is the infinitesimal local neighbourhood of the point x_0^i.

It is possible to consider the parameter x_0^i to be an independent variable, as was proposed by Smith, Paddon [89,90], and turn this approach into a six-dimensional theory of Hyper-Matter. Instead of considering copies of reference frames for all points and the deforming solid as separate entities, Smith, Paddon suggested to introduce a concept of Hyper-Matter – a three dimensional object defined over a six dimensional domain, which includes in itself the deforming body and all the reference frames.

This is an interesting approach, but we do not believe it is particularly useful: it increases the complexity dramatically and does not have any serious advantages, other than the formal convenience of using one single Hyper-Matter object. The solid and the reference frame are aligned locally and only for local operations (such as the calculation of derivatives); considering the displacement field outside the local neighbourhood, even formally, makes no physical sense. Also, all the quantities such as strain, stress and so on, must be converted into six dimensions. Again, this is only a formality that adds complexity without any significant advantages.

5.1.1 Alignment procedure

One last remaining question is to identify exactly what we mean by the alignment of the reference frame and the deforming solid. We need to obtain a rule by which we can determine how to align the reference frame and the solid at any point at any time.

Let us choose and fix a point x_0^i and have our reference frame aligned translationally at this point (see Figure 5.2). We must now determine the procedure for rotational alignment.

Consider the displacement field away from the reference frame within the neighbourhood of the point x_0^i:

$$\mathbf{U} = \mathbf{U}_{x_0^i}\left(x^k, t\right)$$
$$\mathbf{U}\left(x^k, t\right) = U_i\left(x^k, t\right) G^i\left(x^k\right) \tag{5.2}$$

In the last equality the dependency of the displacement field on the point x_0^i, where the reference frame and the solid are aligned, is assumed but not explicitly marked.

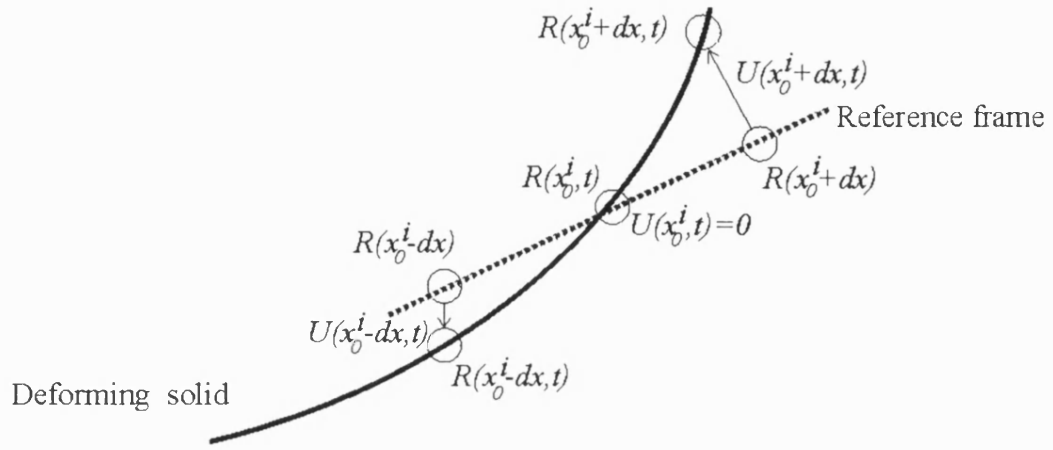


Figure 5.2: Alignment procedure

Let us consider the first derivative of the displacement field $U_i|_j$. As we know this takes part in the expression of the strain of the solid, defined as:

$$\varepsilon_{ij} = \frac{1}{2}\left(U_i|_j + U_j|_i\right) \tag{5.3}$$

Formally we can write an identity:

$$U_i|_j = \frac{1}{2}\left(U_i|_j + U_j|_i\right) + \frac{1}{2}\left(U_i|_j - U_j|_i\right) \tag{5.4}$$

The second order tensor $U_i|_j$ is shown above as a sum of symmetric and antimetric second order tensors. We can see that the strain ε_{ij} is the symmetric part of $U_i|_j$:

$$U_i|_j = \varepsilon_{ij} + \frac{1}{2}\left(U_i|_j - U_j|_i\right) \tag{5.5}$$

We already know that the strain relates to the degree of local deformation suffered by the deforming solid. This result was obtained in section 2.7.2 by considering the change in a small element of length.

Let us now consider the antimetric part of $U_i|_j$:

$$a_{ij} = \frac{1}{2}\left(U_i|_j - U_j|_i\right) \tag{5.6}$$

Since a_{ij} is antimetric, its diagonal components are zero and its off-diagonal components are connected by the relationship $a_{ij} = -a_{ji}$. This means that this tensor has only three independent components, like a vector. We can therefore associate a vector w^k with a_{ij}, with the help of the permutation tensor:

$$-2w^k = \in^{ijk} U_i|_j \tag{5.7}$$

The term $\in^{ijk} U_i|_j$ relates to the curl of the displacement field. Thus, we can rewrite the last equality in a different form in the global Cartesian coordinate system:

$$w^k \mathbf{G}_k = \frac{1}{2} curl\, \mathbf{U} \tag{5.8}$$

From the geometrical meaning of the curl we can conclude that one second of the curl of the displacement field is the average angle of rotation of the deforming body away from the reference frame (see p.86 in Flugge [91] for a detailed discussion). In other words, the vector w^k represents the angle of rotation: its modulus is the actual angle of rotation and its direction is perpendicular to the plane of rotation.

We have arrived at a specific measure of the rotation, expressed through the displacement field. Recalling our expression of the derivative of the displacement field through symmetric and antimetric parts:

$$U_i|_j = \varepsilon_{ij} + \frac{1}{2}\left(U_i|_j - U_j|_i\right) \tag{5.9}$$

we can see that we effectively have a separation into a purely deformational part (the strain tensor ε_{ij}) and a purely rotational part (being the pure rotation of the local neighbourhood of the point).

We can now postulate that the deforming solid and its reference frame are aligned at the point x_0^i, if and only if there is no relative rotation of the deforming body and the reference frame. Quantitatively, this can be expressed in a variety of ways.

The absence of relative rotation means that the angle of rotation $w^k \equiv 0$ at all times. In terms of the displacement field this means that the curl of the displacement field taken at the point x_0^i must be zero at all times:

$$\left(curl\,\mathbf{U}_{x_0^i}\left(x^k,t\right)\right)\Big|_{x_0^i} \equiv 0 \tag{5.10}$$

Alternatively, we can state the deforming body and the reference frame are aligned at a point x_0^i, if and only if the first derivative of the tensor $U_i|_j$ is symmetric:

$$U_i|_j = \varepsilon_{ij} \tag{5.11}$$

Both of these conditions are equivalent and either can be used as required.

In section 6.3, describing the numerical implementation of the method, we will present a numerically fast and simple method of incremental alignments, using the curl of the displacement field.

Having presented the theory of FlexyMatter as an alternative to the global alignment approach, we will now provide its full and formal presentation.

5.2 Formal description of the theory of FlexyMatter

In this section we will formally present the complete theory of FlexyMatter as an extension of the classical theory of elasticity.

As in the classical theory, we have a solid at time *t*, represented by the vector function of the position of the points $\mathbf{R}\left(x^k,t\right)$ in the global coordinate system *K*. We will refer to the solid at the beginning of the experiment as the reference frame $\mathbf{R}\left(x^k\right)=\mathbf{R}\left(x^k,0\right)$, that is, the shape, which we will consider to be the undeformed shape (the undeformed shape is the shape the solid will take if all the applied forces are removed).

It is not required that the undeformed shape is taken as the one the solid had at the start of the experiment ($t = 0$). We can consider the situation where at the start of the experiment the

solid is already deformed. In this case we still assume that $\mathbf{R}(x^k)$ represents the undeformed shape, but, of course, it is no longer equal to the shape of the solid at the start of the experiment: $\mathbf{R}(x^k) \neq \mathbf{R}(x^k,0)$.

As we have identified in the classical theory of elasticity, the following law is universally valid:

$$\ddot{\mathbf{R}}(x^i,t)\rho(x^i,t) = \mathbf{F}(x^i,t) + \mathbf{E}(x^i,t), \forall x^i \in \operatorname{int} I, \forall t > 0 \qquad (5.12)$$

Here $\rho(x^i,t)$ is the material density, $\mathbf{F}(x^i,t)$ is the external volume force and $\mathbf{E}(x^i,t)$ is the internal volume elastic force, which arises due to the deformation of the body away from its original shape and which is expressed in force units per unit volume, and I is the domain where x^i are defined.

In addition to this dynamic condition we also have a static condition at the boundary:

$$\mathbf{S}(x^k,t) = \mathbf{E}_s(x^i,t), \forall x^i \in \partial I \qquad (5.13)$$

This means the applied external boundary force is equal to the reaction elastic boundary force at every boundary point. The boundary elastic force arises due to the deformation of the material away from its original shape and is expressed in force units per unit area.

These two conditions, the dynamic one for the internal points and the static one for the boundary points, are obtained directly from Newton's second law; they are not results from the classical theory of elasticity. The classical theory of elasticity provides a method of evaluating the internal volume elastic force $\mathbf{E}(x^i,t)$ and the boundary elastic force $\mathbf{E}_s(x^i,t)$.

In the classical theory of elasticity, the solid at the start of the experiment is considered a stationary reference frame. The displacement field $\mathbf{U}(x^i,t)$ is defined away from the reference frame to the current position of the solid and all the quantities are expressed in terms of the displacement field. The reference frame is static for all points and at all times.

We propose to use a different mechanism to define the reference frame. A different copy of the reference frame is used for each individual point on the solid; this eliminates the notion of a static reference frame.

We select and fix a point x_0^i on the solid and an instant in time t. In the classical theory of elasticity we would have proceeded by defining the displacement field:

$$\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{R}(x^i) \tag{5.14}$$

Then the strain, the stress and ultimately the elastic forces are defined as the functions of the displacement field expressed in the coordinate system $\mathbf{G}_k(x^i)$:

$$\mathbf{G}_k(x^i)=\frac{\partial \mathbf{R}(x^i)}{\partial x^k} \tag{5.15}$$

However, instead of using the global reference frame $\mathbf{R}(x^i)$, we will use its rotated and translated copy $\mathbf{R}'(x^i)$:

$$\mathbf{R}'(x^i)=\mathbf{A}(t)(\mathbf{R}(x^i)-\mathbf{R}(x_0^i))+\mathbf{R}(x_0^i,t) \tag{5.16}$$

It is rotated around the point $\mathbf{R}(x_0^i,t)$ as its centre of rotation; $\mathbf{A}(t)$ is its matrix of rotation.

The displacement field is defined in the same way, but now using $\mathbf{R}'(x^i)$:

$$\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{R}'(x^i) \tag{5.17}$$

All we now need to do is to provide a condition to determine the rotational matrix $\mathbf{A}(t)$:

$$curl\,\mathbf{U}(x^i,t)\big|_{x_0^i}=0 \tag{5.18}$$

This is the rotational alignment condition. It ensures that the reference frame and the deforming solid are aligned at the point x_0^i on the solid and an instant in time t.

Therefore, comparing our approach to the classical theory of elasticity we can see that we have redefined the reference frame (and the displacement field). Instead of using the global static reference frame, we use its rotated and translated copy for each point and at each instance in time.

From this point, we simply repeat all of the apparatus of the classical theory of elasticity, applied to the above defined displacement field for the current point x_0^i on the solid and at the instant in time t.

5.2.1 Curvilinear coordinate system

We introduce the curvilinear coordinate system $\mathbf{G}_k(x^i)$:

$$\mathbf{G}_k(x^i) = \frac{\partial \mathbf{R}'(x^i)}{\partial x^k} \tag{5.19}$$

In this coordinate system we will express all the quantities, defined at the current point x_0^i.

5.2.2 Strain

We introduce the strain:

$$\varepsilon_{ij} = \frac{1}{2}\left(U_i|_j + U_j|_i\right) \tag{5.20}$$

This is the classical definition of the strain. However, in our case, because of the alignment, the displacement field is such that:

$$curl\,\mathbf{U}(x^i,t)\big|_{x_0^i} = 0 \tag{5.21}$$

Or, in the tensor form:

$$\in^{ijk} U_i|_j (x^p,t)\big|_{x_0^i} = 0 \tag{5.22}$$

Or:

$$U_i|_j (x^p,t)\big|_{x_0^i} = U_j|_i (x^p,t)\big|_{x_0^i} \tag{5.23}$$

Because we will only measure the strain at the point x_0^i, we can write:

$$\varepsilon_{ij} = U_i|_j \tag{5.24}$$

5.2.3 Stress

The stress in the classical theory of elasticity is initially introduced independently from the displacement field (it is only later expressed through it) as an internal elastic force transmitted through an area element. We use the same definition of the stress here.

If we have an area element dA_i, the internal elastic force transmitted through it is $\sigma^{ij} dA_i \mathbf{G}_j$.

Using the same arguments as in the classical theory (the Gauss's Divergence theorem), we can derive an expression for the internal volume elastic force:

$$E^i = \sigma^{ij}\Big|_j \tag{5.25}$$

5.2.4 Constitutive equation

As in the classical theory of elasticity, the stress and the strain are connected by a generic constitutive equation:

$$\sigma^{ij} = E^{ijlm}\varepsilon_{lm} \tag{5.26}$$

For isotropic elastic solids, however, the elastic moduli E^{ijlm} only have two independent parameters. We will use the modulus of elasticity E and the Poisson ratio ν. So far all these considerations and facts are identical to the ones presented in the classical theory of elasticity (refer to chapter 2).

5.2.5 The equations of motion and boundary conditions

The final expression for the internal volume elastic force in the classical theory of elasticity is:

$$E^i = \sigma^{ij}\Big|_j = \frac{E}{2(1+\nu)(1-2\nu)}\left[(1-2\nu)U^i\Big|_j^j + U^j\Big|_j^i\right] \tag{5.27}$$

We can now show that because of the alignment of the reference frame, this expression can be simplified.

Indeed, in our case the displacement field is symmetric:

$$U_i\Big|_j = U_j\Big|_i \tag{5.28}$$

From this we can derive that:

$$U^i\Big|_j^j = U^j\Big|_j^i \tag{5.29}$$

Therefore, finally, the expression for the internal elastic volume force is simply:

$$\sigma^{ij}\Big|_j = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\Big|^j_j \tag{5.30}$$

We can now write the final expression for the elastic force $\mathbf{E}(x^i,t)$:

$$\mathbf{E}(x^i,t) = \sigma^{ij}\Big|_j \mathbf{G}_i(x^k) \tag{5.31}$$

And the final form of the equation of motion of the internal point x_0^i:

$$\ddot{\mathbf{R}}(x_0^k,t)\rho(x_0^k,t) = \mathbf{F}(x_0^k,t) + \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\Big|^j_j (x_0^k,t)\mathbf{G}(x_0^k,t) \tag{5.32}$$

We now proceed to consider the situation when the current point x_0^i is on the boundary of the solid. In the classical theory of elasticity the boundary condition is:

$$\mathbf{S}(x_0^i,t) = \mathbf{E}_s(x_0^i,t) \tag{5.33}$$

Here $\mathbf{E}_s(x_0^i,t)$ is the boundary elastic force transmitted through the boundary point x_0^i.

The boundary elastic force $\mathbf{E}_s(x_0^i,t)$ can be expressed through the stress σ^{ij} as follows:

$$\mathbf{E}_s(x_0^k,t) = \sigma^{ij}(x_0^k,t)n_j(x_0^k,t)\mathbf{G}_i(x_0^k) \tag{5.34}$$

The vector n_j here is the outward normal to the boundary surface at the point x_0^i.

Or, in full, the expression for the elastic boundary force can written as follows:

$$\mathbf{E}_s(x_0^k,t) = \frac{E}{1+\nu}\left(U^i\Big|^j(x^k,t) + \frac{\nu}{1-2\nu} U^m\Big|_m (x^k,t)G^{ij}(x^k)\right) n_j(x_0^k,t)\mathbf{G}_i(x_0^k) \tag{5.35}$$

This is the boundary condition for the current boundary point x_0^i.

We can now provide the final equations of motion and the boundary condition for all points in the solid:

$$\begin{cases} \ddot{\mathbf{R}}(x^k,t)\rho(x^k,t) = \mathbf{F}(x^k,t) + \dfrac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\big|^j_j(x^k,t)\mathbf{G}_i(x^k,t), \\ \forall x^i \in \operatorname{int} I, \forall t > 0 \\ \mathbf{S}(x^i,t) = \dfrac{E}{1+\nu}\left(U^i\big|^j(x^k,t) + \dfrac{\nu}{1-2\nu} U^m\big|_m(x^k,t) G^{ij}(x^k) \right) n_j(x^k,t)\mathbf{G}_i(x^k), \\ \forall x^i \in \partial I, \forall t > 0 \end{cases} \tag{5.36}$$

The displacement field here is assumed to have been calculated as described in Chapter 5: for each individual point and each instance in time, its own reference frame is aligned rotationally and transitionally. Then the displacement field is calculated for that point away from its individual reference frame. All the derivatives of the displacement field are then calculated as normal and their values are taken at the current point.

5.2.6 The parameters of the process of calculation of the displacement field

In this section we have intentionally omitted any special notation (other than using a simple prime) to reflect the fact that a different reference frame is used for each individual point. It therefore formally depends both on the point for which it is calculated and the instance in time when it is calculated (we will call these the alignment point and the current time). It is clear that once the reference frame is dependent on the point as well as time all the other quantities (displacement field, strain, stress and elastic forces) will also depend on which point and at what time they were calculated.

We decided not to formally identify this dependency to avoid unnecessary complication of the theory. The point for which the reference frame is aligned and the time at which it happens are parameters of the process, and therefore of all the quantities which are calculated during the process. These parameters are assumed to be fixed during the calculation of the derivatives at a given point, as well as all other quantities that depend on the derivatives, such as displacement field, strain, stress and elastic forces.

Smith, Paddon [89,90] formally incorporated this dependency into the notation for all these quantities defined at a point. However, including these parameters formally into the notation turns the extended classical theory of elasticity into a new six-dimensional theory, requiring a completely different presentation. All the quantities defined at a point on the solid must then be viewed as defined over a six-dimensional domain. All the operations on these six-dimensional quantities must be redefined (or at least explicitly explained), including

derivation and integration. New operations (such as taking a collateral derivative) must be added.

We believe this complexity does not justify itself. Presenting this extension to the classical theory of elasticity as a full-blown six-dimensional theory of a six-dimensional deforming Hyper-Matter is not justified. Formally speaking, in the six-dimensional domain, we only ever need to consider the three-dimensional diagonal and its infinitesimally thin six-dimensional neighbourhood (see a simple illustration Figure 5.3 below). Anything outside this infinitesimally thin six-dimensional neighbourhood of the three-dimensional diagonal is of no interest.

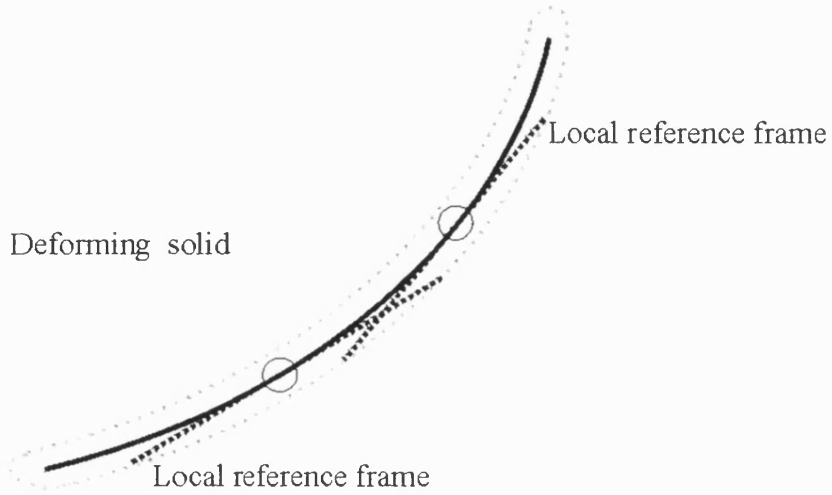


Figure 5.3: Six-dimensional solid and its infinitesimal neighbourhood

Another approach, adopted in this thesis, is to keep the alignment point and the current time as implicit parameters, as the part of the process of defining the displacement field in the neighbourhood of a point (recognising the fact that we will only ever need either first or second derivative of the displacement field). We believe this approach is advantageous, because it keeps the theory simple and close to the classical theory of elasticity, whilst only adding as much complexity as absolutely necessary.

5.2.7 Calculation of the first and second derivative of the displacement field

We have already presented one simplification in the calculation of the internal volume elastic force arising from the fact that the reference frame is aligned specifically for each individual point. We will now show there is another simplification due to the translational

alignment: the first covariant derivative of the displacement field can be calculated as an ordinary partial derivative.

Let us consider a point x_0^i. The fact that the reference frame for this point is aligned both translationally and rotationally with the deforming solid, can be expressed via the displacement field in the following way:

$$\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{R}'(x^i) \tag{5.37}$$

Where the reference frame is:

$$\mathbf{R}'(x^i)=\mathbf{A}(t)(\mathbf{R}(x^i)-\mathbf{R}(x_0^i))+\mathbf{R}(x_0^i,t) \tag{5.38}$$

Combining these two expressions into one, we obtain:

$$\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{A}(t)(\mathbf{R}(x^i)-\mathbf{R}(x_0^i))-\mathbf{R}(x_0^i,t) \tag{5.39}$$

Taking the value of the displacement field at the point x_0^i itself, we can see that it is zero:

$$\mathbf{U}(x_0^i,t)=0 \tag{5.40}$$

This is true for any point x_0^i and any time t.

In coordinate form, this simply means that each of the coordinates of the displacement field is zero:

$$U_i(x_0^k,t)=0 \tag{5.41}$$

Let us now look at the expression for the first covariant derivative of the displacement field:

$$U_i|_j = U_{i,j} - U_s\Gamma_{ij}^s \tag{5.42}$$

This is the classical expression for the first covariant derivative via the Christoffel symbols Γ_{ij}^s, which describe the topology of the underlying coordinate system.

In our case, because $U_i(x_0^k,t)=0$, we can see that:

$$U_i|_j(x_0^k,t)=U_{i,j}(x_0^k,t) \tag{5.43}$$

We can use the normal partial derivative instead of the full covariant derivative to calculate the first derivative of the displacement field.

We will now show that it is generally impossible, however, to use the second partial derivative instead of the second covariant derivative.

Indeed, the formal expression of the second derivative is:

$$U_i\big|_{jk} = U_i\big|_{j,k} - U_s\big|_j \Gamma^s_{ik} - U_i\big|_s \Gamma^s_{kj} \Leftrightarrow$$
$$U_i\big|_{jk} = U_{i,jk} - U_{s,j}\Gamma^s_{ik} - U_{i,s}\Gamma^s_{kj} \tag{5.44}$$

We can see that using the second partial derivative instead of the full covariant one is equivalent to dropping the last two terms on the right. These two terms are generally non-zero (unless in the Cartesian coordinate system) and dropping them will create an error. As we know from the expression for the internal volume elastic force, the second derivative of the displacement field is a force-like quantity.

The error created by dropping these two terms will produce an error in the calculation of the elastic force, which is the more pronounced the more curvilinear the coordinate system on the reference frame is. It is possible to come up with a number of examples, where this will lead to appreciable errors and seriously wrong simulation results. Figure 5.4 shows an illustrative example of the kind of errors that can be encountered.

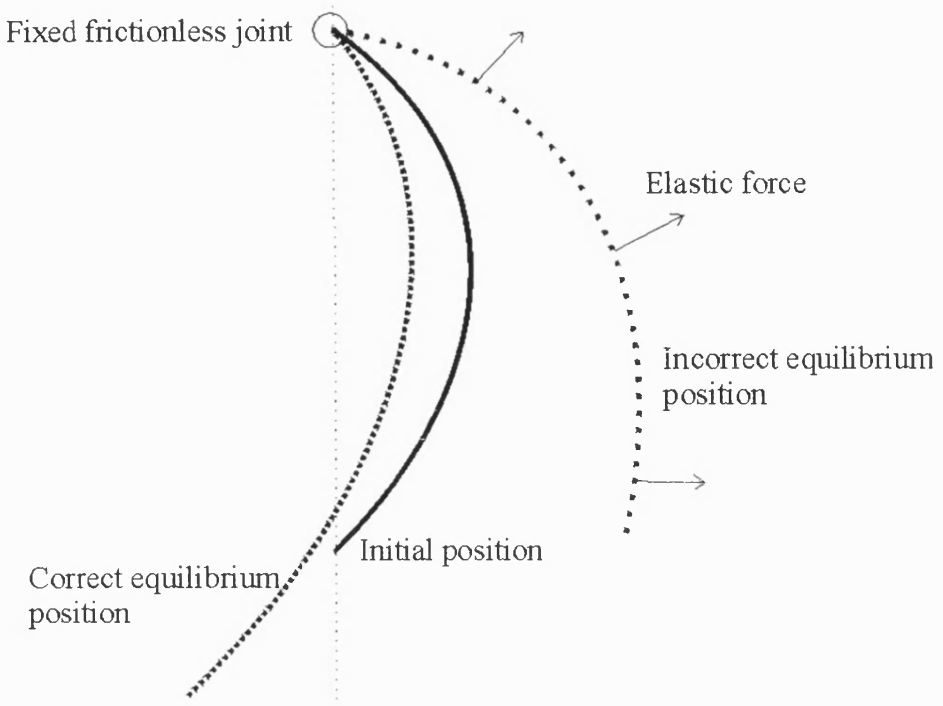


Figure 5.4: Example of an incorrect evaluation of the elastic force due to using second partial derivative instead of the second covariant derivative

In this example, a curved, thin and flexible solid is suspended in a gravitational field at one end. The suspension point is a frictionless joint. The solid should assume its correct equilibrium position, as shown on the picture. Instead, due to the erroneous evaluation of the elastic force, it assumes an incorrect equilibrium position, very different from the correct one.

5.3 Summary

In this chapter, we gave a complete presentation of our theory of FlexyMatter, as the extension of the linear classical theory of elasticity with local rotational alignments of the deforming solid and the reference frame.

We started with an informal introduction in section 5.1. We introduced the notion of local alignment of the reference frame and the deforming body and then described in detail the alignment procedure in section 5.1.1.

In section 5.2 we formally presented our theory of FlexyMatter. Because it is an extension of the linear classical theory of elasticity, this presentation relates heavily to the classical theory presented in chapter 2 and is therefore relatively brief. We introduced the local alignment of the deforming solid and the reference frame as the additional first step in the classical process of calculation of the elastic force at a point. Then the displacement field

can be defined, followed by all the other classical quantities: strain, stress and finally the elastic force. We followed with the presentation of the constitutive equations in section 5.2.4 and the equations of motion in section 5.2.5. Finally, in the last section 5.2.7 we presented simplifications in the classical framework, due to the local alignments.

In the next chapter 6, we present a practical numerical solution of the equations of motion presented in section 5.2.5.

6. Practical solution of the modelling problem

In this chapter we will consider practical methods of solving the modelling problem we have described in detail in the previous chapter. First of all, let us recap the formal statement of the modelling problem.

We have a solid, which is given at the beginning of the experiment by a mapping function:

$$\mathbf{R}(x^i): I \to B \tag{6.1}$$

Where the domain I is defined as:

$$\begin{aligned} &I = I[A^1, A^2, A^3] = I^3[A^1, A^2, A^3] = \\ &= \left\{(x^1, x^2, x^3): x^1 \in [0, A^1], x^2 \in [0, A^2], x^3 \in [0, A^3]\right\} \end{aligned} \tag{6.2}$$

and the set B is the image of I under the mapping $\mathbf{R}(x^i)$.

We also have the following quantities:

- A volume force field $\mathbf{F}$ at every internal point x^k and at any time t:

$$\mathbf{F} = \mathbf{F}(x^k, t), \forall x^i \in \operatorname{int} I, t \geq 0 \tag{6.3}$$

- A surface force field $\mathbf{S}$ at every boundary point x^k and at any time t:

$$\mathbf{S} = \mathbf{S}(x^i, t), \forall x^i \in \partial I, t \geq 0 \tag{6.4}$$

- Material density ρ at every internal point x^k and at any time t:

$$\rho = \rho(x^i,t), \forall t \geq 0, \forall x^i \in \partial I \tag{6.5}$$

- The initial velocity $\mathbf{V}$:

$$\mathbf{V} = \mathbf{V}(x^i) = \dot{\mathbf{R}}(x^i), \forall x^i \in I \tag{6.6}$$

These are the initial conditions, quantities that are known before the start of the experiment.

We have also identified earlier in this chapter that, during the experiment, the evolution of the solid subject to the above-given initial conditions and the applied forces will be described by the following set of dynamic and static differential equations:

$$\begin{cases} \ddot{\mathbf{R}}(x^k,t)\rho(x^k,t) = \mathbf{F}(x^k,t) + \dfrac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\big|^j_j(x^k,t)\mathbf{G}_i(x^k), \\ \forall x^i \in \operatorname{int} I, \forall t > 0 \\ \mathbf{S}(x^i,t) = \dfrac{E}{1+\nu}\left(U^i\big|^j(x^k,t) + \dfrac{\nu}{1-2\nu} U^m\big|_m(x^k,t) G^{ij}(x^k) \right) n_j(x^k,t)\mathbf{G}_i(x^k), \\ \forall x^i \in \partial I, \forall t > 0 \\ \mathbf{R}(x^k,0) = \mathbf{R}(x^k), \forall x^i \in I \\ \dot{\mathbf{R}}(x^k,0) = \mathbf{V}(x^k), \forall x^i \in I \end{cases} \tag{6.7}$$

Where: $\mathbf{R}(x^i,t): I \to B(t), t > 0$ is the mapping function for the deforming solid $B(t)$, E is the modulus of elasticity, ν is the Poisson ratio (these two scalar constants describe the material properties of the solid), $U(x^i,t)$ is the displacement field and $\mathbf{G}_i(x^k,t)$ are the covariant base vectors defined as:

$$\mathbf{G}_i(x^k,t) = \frac{\partial \mathbf{R}(x^k,t)}{\partial x^i}, \forall x^i \in I, \forall t \geq 0 \tag{6.8}$$

We need to clarify the definition of the derivative at the boundary points. We do it in the usual way, having defined the covariant base vectors at the internal points $\forall x^i \in \operatorname{int} I$, we simply extend them continuously to the boundary:

$$\mathbf{G}_i(x^k,t) = \lim_{\hat{x}^k \to x^k} \mathbf{G}_i(\hat{x}^k,t) = \lim_{\hat{x}^k \to x^k} \frac{\partial \mathbf{R}(\hat{x}^k,t)}{\partial x^i}, \forall x^i \in \partial I, \forall \hat{x}^i \in \operatorname{int} I, \forall t \geq 0 \tag{6.9}$$

The last two equations in (6.7) are the initial conditions for the first and second equations, specifying the initial position and the initial velocity of the solid.

The displacement field $U(x^i,t)$ is defined as:

$$\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{R}'(x^i) \tag{6.10}$$

Here, $\mathbf{R}'(x^i)$ is the aligned reference frame (which is the solid at the start of the experiment $\mathbf{R}(x^i)$). The following alignment procedure for the current point x_0^k and the current instance in time t_0 is used:

- First, we align the reference frame and the deforming solid translationally, that is, we make sure that:

$$\mathbf{R}'(x_0^i)=\mathbf{R}(x_0^i,t_0) \tag{6.11}$$

- Secondly, we align the reference frame and the deforming solid rotationally, so that the deformation field $\mathbf{U}(x^i,t)=\mathbf{R}(x^i,t)-\mathbf{R}'(x^i)$, calculated away from the reference frame satisfies the following condition:

$$(curl\,\mathbf{U}(x^i,t))\big|_{x^i=x_0^i,t=t_0}=0 \tag{6.12}$$

Or, in tensor form:

$$\in^{jil} U_i|_j(x_0^k,t_0)=0 \Leftrightarrow U_i|_j(x_0^k,t_0)=U_j|_i(x_0^k,t_0) \tag{6.13}$$

As we have noted before, this is the same as:

$$U_{i,j}(x_0^k,t_0)=U_{j,i}(x_0^k,t_0) \tag{6.14}$$

From the point of view of the physical experiment, we want to find the shape $B(t)$ of the solid B after the start of the experiment, which is subject to the applied forces and initial conditions. Mathematically speaking, it is equivalent to the problem of finding a vector function:

$$\begin{gathered}\mathbf{R}(x^i,t): I \to B(t), t \geq 0 \\ \mathbf{R}(x^i,0)=\mathbf{R}(x^i)\end{gathered} \tag{6.15}$$

which will satisfy the above-given system of differential equations (6.7) for all points $x^k \in I$ and for all times $t \geq 0$. This is the ideal situation since, having obtained the solution $\mathbf{R}(x^i,t)$, we could use it to obtain the positions of any point at any time.

In practice, however, it is not a simple task due to a number of reasons:

- The solid in its initial position B is rarely given by a continuous and differentiable mapping $\mathbf{R}(x^i)$. Very often, all we have is a set of points in the global coordinate system representing the body.
- Even if we had a continuous and differentiable mapping $\mathbf{R}(x^i)$ giving the solid at its initial position, it is in general impossible to solve the system of equations (6.7) explicitly for any $\mathbf{F}(x^k,t)$, $\mathbf{S}(x^i,t)$ and the initial conditions $\mathbf{R}(x^k)$, $\mathbf{V}(x^k)$ to obtain $\mathbf{R}(x^i,t)$ (unless in very simple cases).

The only practically feasible way to obtain a general solution is to convert the continuous problem (6.7) into a discrete problem, which approximates our continuous problem. This approximate discrete problem should have its approximation error tending to zero as the number of discretization nodes tends to infinity.

6.1 Derivation of an approximate discrete model for our precise problem

Our aim in this section is to derive an approximate discrete model for our precise continuous model.

Our precise continuous model M operates in terms of continuous quantities: the body is a continuous set of points; the external force $\mathbf{F}(x^k,t)$ is a continuous (as well as differentiable) vector function and so on. As we have noted earlier, this is the only way we could apply the whole of the necessary mathematical apparatus to the original physical problem of modelling real physical solids, deforming under the action of applied forces and find a solution.

We have converted the physical problem into a mathematical form, so instead of the real physical body we considered an idealised body B, introduced a notion of a mapping function $\mathbf{R}(x^k)$ and the velocity $\mathbf{V}(x^k)$. We said that the external forces $\mathbf{F}$ and $\mathbf{S}$ will also be represented by continuous functions $\mathbf{F}(x^k,t)$ and $\mathbf{S}(x^i,t)$ and the mass of the body will be represented via a continuous scalar density function $\rho(x^k,t)$. We have also introduced a vector function $\mathbf{R}(x^i,t)$, which represents the deforming body throughout the experiment. Finally, we formally identified that to solve the modelling problem we must find this vector function $\mathbf{R}(x^i,t)$.

Having converted the physical problem into the required mathematical form, we then followed a classical path for deriving the solution:

- We used Newton's second law to derive that $\mathbf{R}(x^i,t)$ satisfies a certain system of equations: $\ddot{\mathbf{R}}(x^k,t)\rho(x^k,t)=\mathbf{F}(x^k,t)+\mathbf{E}(x^k,t)$ for the internal points and $\mathbf{S}(x^i,t)=\mathbf{E}_s(x^k,t)$ for the boundary points, where $\mathbf{E}(x^k,t)$ was called the internal volume elastic force and $\mathbf{E}_s(x^k,t)$ - the boundary elastic force.

  We have therefore identified that to find $\mathbf{R}(x^i,t)$ we need to find the elastic forces $\mathbf{E}(x^k,t)$ and $\mathbf{E}_s(x^k,t)$.

- To find the elastic forces $\mathbf{E}(x^k,t)$ and $\mathbf{E}_s(x^k,t)$, we have introduced notions of the displacement field $\mathbf{U}(x^i,t)$, the strain tensor $\varepsilon_{ij}(x^k,t)$ and the stress tensor $\sigma^{ij}(x^k,t)$.

- Then, based on a number of practical observations, we derived the expressions for the elastic forces $\mathbf{E}(x^k,t)$ and $\mathbf{E}_s(x^k,t)$ via the displacement field $\mathbf{U}(x^i,t)$ (which itself is expressed via the reference frame and the deforming body aligned at each point).

Finally, we have obtained a system of equations (6.7) for the unknown vector function $\mathbf{R}(x^i,t)$. The system is a set of dynamic and static differential equations for $\mathbf{R}(x^i,t)$.

Theoretically speaking, we only have one step remaining - to solve the system of equations (6.7) and obtain explicit expressions for $\mathbf{R}(x^i,t)$. But, as we have mentioned in the previous section, this is generally an impossible task. There is no known general technique to solve (6.7) in its continuous form. We are therefore forced to look for ways to bypass this difficulty.

The most obvious way is to discretize our continuous modelling problem M and consider a discrete model $M[D]$, where D is a positive whole number, relating to the degree of discretization, so that $M[D]$ tends to a continuous model M' when D tends to infinity.

We want $M[D]$ to approximate our continuous model M:

$$M[D] \underset{D\to\infty}{\to} M$$

To obtain our approximation $M[D]$, we need to perform the following tasks:

- Turn all our continuous quantities into discrete, that is, turn functions defined over continuous domains into discrete ones defined over domains consisting of discrete sets of points.
- Discretize the set of equations (6.7), which will entail defining discrete analogues for all the operations used in (6.7) (mostly differentiation).

In this scenario, the parameter D will be physically the number of nodes where all the functional quantities are defined.

We will now proceed to describe a practical method to obtain the approximating model $M[D]$. We will do this in two stages: firstly, we introduce the spatial discretization and secondly we introduce the discretization in time (this will include the numerical solution method for stepping through time).

6.1.1 The spatial discretization of the model

The spatial discretization will consist of the discretization of the domain I and therefore all the functional quantities as well as the introduction of discrete analogues to the continuous spatial operations in (6.7) (differentiation).

We recall that all the functional quantities, that the continuous model M is dealing with, are defined over a continuous domain $I = I[A^1, A^2, A^3]$. As the first step we therefore discretize the domain I. We split the interval $[0, A^1]$ into $L+1$ nodes, the interval $[0, A^2]$ into $M+1$ nodes and the interval $[0, A^3]$ into $N+1$ nodes (see Figure 6.1 below).

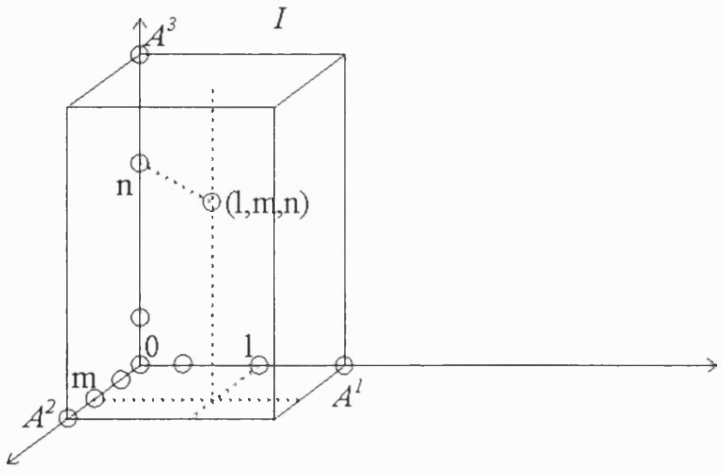


Figure 6.1: Discretization of the continuous domain I

Each interval is split uniformly:

- The interval $\left[0, A^1\right]$ into the nodes $\left\{\frac{lA^1}{L}, 0 \le l \le L\right\}$.
- The interval $\left[0, A^2\right]$ into the nodes $\left\{\frac{mA^2}{M}, 0 \le m \le M\right\}$.
- The interval $\left[0, A^3\right]$ into the nodes $\left\{\frac{nA^3}{N}, 0 \le n \le N\right\}$.

The corresponding distances between the nodes for each interval are: $\Delta l = \frac{A^1}{L}$, $\Delta m = \frac{A^2}{M}$ and $\Delta n = \frac{A^3}{N}$.

We have therefore obtained a discrete approximation $I[D]$ for the continuous domain I:

$$I[D] = \left\{\left(\frac{lA^1}{L}, \frac{mA^2}{M}, \frac{nA^3}{N}\right), 0 \le l \le L, 0 \le m \le M, 0 \le n \le N\right\} \tag{6.16}$$

The discretization parameter D is the total number of all discrete nodes in $I[D]$:

$$D = (L+1)(M+1)(N+1) \tag{6.17}$$

We can refer to the individual points in $I[D]$ as (l, m, n).

We can now consider all continuous functional quantities, defined over I, over this discrete domain $I[D]$. Thus we will obtain discretized versions of all our continuous quantities:

$$\begin{aligned} \mathbf{R}(l,m,n) &= \mathbf{R}\left(\frac{lA^1}{L}, \frac{mA^2}{M}, \frac{nA^3}{N}\right), 0 \le l \le L, 0 \le m \le M, 0 \le n \le N \\ \mathbf{R}(l,m,n,t) &= \mathbf{R}\left(\frac{lA^1}{L}, \frac{mA^2}{M}, \frac{nA^3}{N}, t\right), 0 \le l \le L, 0 \le m \le M, 0 \le n \le N \end{aligned} \tag{6.18}$$

and so on for all other quantities.

We will now define discrete analogues to the quantities, used in the system of equations (6.7).

We start with the covariant base vectors $\mathbf{G}_i(x^k)$:

$$\mathbf{G}_1(l,m,n)=\begin{cases}\dfrac{\mathbf{R}(l+1,m,n)-\mathbf{R}(l,m,n)}{\Delta l}, 0\le l<L\\ \dfrac{\mathbf{R}(l,m,n)-\mathbf{R}(l-1,m,n)}{\Delta l}, l=L\end{cases} \tag{6.19}$$

Similar expressions are defined for the other two discrete versions of the base vectors $\mathbf{G}_2(l,m,n)$ and $\mathbf{G}_3(l,m,n)$. Having defined $\mathbf{G}_i(l,m,n)$, we also have defined the metric tensor in its covariant form:

$$G_{ij}(l,m,n)=\mathbf{G}_i(l,m,n)\cdot\mathbf{G}_j(l,m,n) \tag{6.20}$$

We will also need the contravariant basis $\mathbf{G}^i(l,m,n)$ and the contravariant form of the metric tensor $G^{ij}(l,m,n)$. The following expression is true for the covariant and contravariant base vectors:

$$\in_{ijk}\mathbf{G}^i=\mathbf{G}_j\times\mathbf{G}_k \tag{6.21}$$

If we write it for a specific case, say $i=1,j=2,k=3$, we obtain a simple equality expressing the contravariant base vector $\mathbf{G}^1$ via the covariant base vectors $\mathbf{G}_2$ and $\mathbf{G}_3$:

$$\in_{123}\mathbf{G}^1=\mathbf{G}_2\times\mathbf{G}_3\Rightarrow\mathbf{G}^1=\frac{1}{\sqrt{G}}\mathbf{G}_2\times\mathbf{G}_3=\frac{1}{\sqrt{\mathbf{G}_1\cdot(\mathbf{G}_2\times\mathbf{G}_3)}}\mathbf{G}_2\times\mathbf{G}_3 \tag{6.22}$$

Similar expressions can be obtained for the other two contravariant base vectors $\mathbf{G}^2$ and $\mathbf{G}^3$.

The metric tensor in contravariant form is then:

$$G^{ij}(l,m,n)=\mathbf{G}^i(l,m,n)\cdot\mathbf{G}^j(l,m,n) \tag{6.23}$$

These base vectors are defined for the undeformed body in its original position. The ones that are used in (6.7) are aligned versions of them. We align the base vectors $\mathbf{G}_i(l,m,n)$ and $\mathbf{G}^i(l,m,n)$ with the deforming body at each point to remove the local rotational components of the displacement field, which, as we have shown in section 4.2.1, introduce serious errors in the evaluation of the elastic forces.

The alignment procedure does not change the vectors $\mathbf{G}_i(l,m,n)$ or $\mathbf{G}^i(l,m,n)$ themselves; the ones that are used in (6.7) are simply rotated copies of the original $\mathbf{G}_i(l,m,n)$ and $\mathbf{G}^i(l,m,n)$. There is a definite algorithm for aligning the base vectors for each point at each moment in time. However, to introduce the algorithm we need the discretization in time, therefore we will postpone the formal description of the alignment procedure until the discretization in time has been discussed. For the time being, we simply assume that both sets $\mathbf{G}_i(l,m,n)$ and $\mathbf{G}^i(l,m,n)$ are aligned for each point and at all times as required.

We can formally write the discrete version of (6.7). This will help us see what we have so far and what still needs to be done:

$$\begin{cases} \ddot{\mathbf{R}}(l,m,n,t)\rho(l,m,n,t) = \mathbf{F}(l,m,n,t) + \dfrac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\big|_j^j(l,m,n,t)\mathbf{G}_i(l,m,n) \\ \mathbf{S}(l,m,n,t) = \\ = \dfrac{E}{1+\nu}\left(U^i\big|^j(l,m,n,t) + \dfrac{\nu}{1-2\nu} U^m\big|_m(l,m,n,t) G^{ij}(l,m,n) \right) n_j(l,m,n,t)\mathbf{G}_i(l,m,n) \\ \text{where } 0 < l < L, 0 < m < M, 0 < n < N \\ \mathbf{R}(l,m,n,0) = \mathbf{R}(l,m,n) \\ \dot{\mathbf{R}}(l,m,n,0) = \mathbf{V}(l,m,n) \\ \text{where } l = 0 \vee l = L \wedge m = 0 \vee m = M \wedge n = 0 \vee n = N \end{cases} \tag{6.24}$$

The following quantities still need to be formally defined: $U^i\big|_j^j(l,m,n,t)$, $U^i\big|^j(l,m,n,t)$, $U^m\big|_m(l,m,n,t)$ and $n_j(l,m,n,t)$.

The vector $n_j(l,m,n,t)$ is the unit normal vector at a boundary point (l,m,n), pointing outwards, resolved in the contravariant basis $\mathbf{G}^i(l,m,n)$.

Let us now consider the displacement field $\mathbf{U}(l,m,n,t)$. As we know it is defined as:

$$\mathbf{U}(l',m',n',t) = \mathbf{R}(l',m',n',t) - \mathbf{R}'(l,m,n) \tag{6.25}$$

Here $\mathbf{R}'(l,m,n)$ is the aligned copy of the reference frame for the point (l,m,n) (we will consider the alignment procedure later).

We can resolve the displacement field into its contravariant components in the covariant basis of the point (l,m,n), as usual:

$$\mathbf{U}(l',m',n',t)=U^i(l',m',n',t)\mathbf{G}_i(l,m,n) \tag{6.26}$$

We need to define various derivatives of the displacement field, taken at the centre point (l,m,n). For this we will only need the immediate neighbourhood of the point $(l',m',n')\in\{(l\pm1,m\pm1,n\pm1)\}$ (obvious range restrictions apply for boundary points).

We will now consider the first derivative of the displacement field:

$$U^i\big|_j(l,m,n,t)=\left(U^i\big|_j(l',m',n',t)\right)\Big|_{\substack{l'=l\\m'=m\\n'=n}}=U^i{}_{,j}(l,m,n,t) \tag{6.27}$$

The fact that the first covariant derivative of the displacement field actually equals the partial derivative was discussed in section 5..

The partial derivative is approximated as usual:

$$U^i{}_{,1}(l,m,n,t)=\begin{cases}\dfrac{U^i(l+1,m,n)-U^i(l,m,n)}{\Delta l}=\dfrac{U^i(l+1,m,n)}{\Delta l}, & 0\le l<L\\[2ex] \dfrac{U^i(l,m,n)-U^i(l-1,m,n)}{\Delta l}=\dfrac{-U^i(l-1,m,n)}{\Delta l}, & l=L\end{cases} \tag{6.28}$$

And similarly for $U^i{}_{,2}(l,m,n,t)$ and $U^i{}_{,3}(l,m,n,t)$.

We can now write down the expressions for the terms $U^m\big|_m(l,m,n,t)$ and $U^i\big|^j(l,m,n,t)$:

$$U^m\big|_m(l,m,n,t)=U^1{}_{,1}(l,m,n,t)+U^2{}_{,2}(l,m,n,t)+U^3{}_{,3}(l,m,n,t)$$

$$U^i\big|^j(l,m,n,t)=U^i{}_{,s}(l,m,n,t)G^{sj}(l,m,n)= \tag{6.29}$$

$$=U^i{}_{,1}(l,m,n,t)G^{1j}(l,m,n)+U^i{}_{,2}(l,m,n,t)G^{2j}(l,m,n)+U^i{}_{,3}(l,m,n,t)G^{3j}(l,m,n)$$

We will now consider the last term $U^i\big|^j_j(l,m,n,t)$, involving the second derivative.

First of all we recall that the expression for the second derivative of the displacement field is:

$$U^i\big|_{jk}(l,m,n,t)=U^i{}_{,jk}(l,m,n,t)+U^l{}_{,j}(l,m,n,t)\Gamma^i_{kl}(l,m,n,t)+U^i{}_{,l}(l,m,n,t)\Gamma^l_{jk}(l,m,n,t) \quad (6.30)$$

In the expression on the right we have only explicitly defined the first derivative $U^i{}_{,j}$. All the other terms are still to be defined.

The second partial derivative of the displacement field is as usual:

$$U^i{}_{,11}(l,m,n,t)=\frac{U^i(l+1,m,n,t)-2U^i(l,m,n,t)+U^i(l-1,m,n,t)}{(\Delta l)^2}=$$

$$=\frac{U^i(l+1,m,n,t)+U^i(l-1,m,n,t)}{(\Delta l)^2} \quad (6.31)$$

$$U^i{}_{,12}(l,m,n,t)=\frac{U^i(l+1,m+1,n,t)-U^i(l,m+1,n,t)-U^i(l+1,m,n,t)+U^i(l,m,n,t)}{(\Delta l)(\Delta m)}=$$

$$=\frac{U^i(l+1,m+1,n,t)-U^i(l,m+1,n,t)-U^i(l+1,m,n,t)}{(\Delta l)(\Delta m)}$$

Similar expressions can be derived for all other combinations of indices i and j.

We now need to derive approximate expressions for the Christoffel symbols $\Gamma^i_{jk}(l,m,n,t)$. One of the expressions involving these Christoffel symbols is:

$$\mathbf{G}^i_{,j}(l,m,n,t)=-\Gamma^i_{jk}(l,m,n,t)\mathbf{G}^k(l,m,n) \quad (6.32)$$

The approximate expression for the first derivative of the contravariant base vector is:

$$\mathbf{G}^i_{,1}(l,m,n,t)=\frac{\mathbf{G}^i(l,m,n,t)-\mathbf{G}^i(l-1,m,n,t)}{\Delta l} \quad (6.33)$$

And similarly for other values of j.

Now, to obtain the Christoffel symbols, we must resolve this vector into contravariant components:

$$\mathbf{G}^i_{,j}(l,m,n,t)=A^i_{jk}\mathbf{G}^k(l,m,n) \quad (6.34)$$

Christoffel symbols will be expressed through the coefficients of this decomposition:

$$\Gamma^i_{jk} = -A^i_{jk} \tag{6.35}$$

We can now proceed to obtain the final expression for the term $U^i\big|^j_j(l,m,n,t)$:

$$U^i\big|^j_j(l,m,n,t) = U^i\big|_{jk}(l,m,n,t)G^{kj}(l,m,n,t) \tag{6.36}$$

This concludes the formal stage of spatial discretization. We have defined all the quantities taking part in the approximate system of equations (6.24).

The next stage will be discuss the discretization in time and consider in detail the procedure for the dynamic alignment of the reference frame and the deforming solid.

However, before we move on, we need to show that the spatial discretization we have described here will produce an approximate model in its true sense. That is, we will answer the question: if we start to increase the number of nodes (parameter D) and send it to infinity (uniformly, in all three indices L, M, and N), will we get in the limit our continuous model M?

The answer to this question is, of course, yes. Due to the way we have discretized the domain I, our uniform mesh of nodes does indeed cover the whole of I in the limit:

$$I[D] \underset{\substack{L\to\infty \\ M\to\infty \\ N\to\infty}}{\longrightarrow} I \tag{6.37}$$

We will now check if all the approximate expressions for the differentiation operations we have introduced will tend to the real derivatives in the limit. It is indeed so because we have used the classical definitions of the derivatives. If we consider, for example, the first derivative of the displacement field:

$$U^i{}_{,1}(l,m,n,t) = \begin{cases} \dfrac{U^i(l+1,m,n)-U^i(l,m,n)}{\Delta l} = \dfrac{U^i(l+1,m,n)}{\Delta l}, 0 \le l < L \\ \dfrac{U^i(l,m,n)-U^i(l-1,m,n)}{\Delta l} = \dfrac{-U^i(l-1,m,n)}{\Delta l}, l = L \end{cases} \tag{6.38}$$

This is a classical definition of the first partial derivative of a scalar function U^i. $U^i(l,m,n)$ is always zero for any point at any time due to our translational alignment. As $L \to \infty$ the distance Δl between the nodes of the mesh will tend to zero and so will $U^i(l+1,m,n)$ (or

$-U^i(l-1,m,n)$ depending on the position of the point). Their ratio will tend to the first derivative of the U^i by x^1 if it exists (we assume it does).

The same is true for all other approximate expressions for the first and second derivatives of the displacement field taking part in (6.24).

6.1.2 Discretization in time

Before we consider the problem of discretization in time and the associated numerical solution techniques, let us examine the approximate system of equations (6.24) in detail.

We have so far completed the spatial discretization of the model. We turned the continuous domain I into a discrete domain $I[D]$ and obtained discrete versions of all the terms in (6.24), defined at discrete nodes; we have not made any significant changes to the system. In (6.24) we have, in principle, three types of equations:

- Dynamic equations for the internal points (the first set of equations in (6.24)). The second derivative of $\mathbf{R}$ in time (the acceleration of the material point) takes part in these equations.
- Static equations for the boundary points (the second set of equations in (6.24)). These equations describe the balance of forces transmitted through the boundary of the solid.
- Initial conditions. These conditions describe the initial position and velocity of all the points of the solid at the start of the experiment.

In principal, any numerical solution technique for ordinary differential equations deals with the system of the following form:

$$\begin{cases} \dot{\mathbf{r}}(t) = \mathbf{f}(t, \mathbf{r}(t)), t > 0 \\ \mathbf{r}(0) = \mathbf{r}_0 \end{cases} \tag{6.39}$$

Our system (6.24) differs from this form. The fact that we have the second derivative in time is of no importance (it can be easily converted into a system with only one derivative in time). The important problem is with our static equations: we have a system of nodes belonging to the solid; for some of these nodes (the internal ones) we have a dynamic system of equations and for the other nodes (the boundary ones) we have a static system of

equations. More importantly, these systems of equations are coupled and so must be solved together.

For the continuous (idealised) solid this is an important feature of the system describing its evolution: the situation is different in principal for internal points as opposed to the boundary points. An internal point has density and volume associated with it and therefore has acceleration, while a boundary point does not have density or mass, only boundary surface area that can be associated with it. Because of this, we have to deal with two different systems of equations describing the behaviours of the internal and boundary points of the solid. The dynamic equations, combining the acceleration, external volume force and the internal elastic volume force, describe the evolution of the internal points, while the static equations, describing the balance of the external and internal elastic forces, transmitted through elements of the solid's boundary, are used to implicitly describe the evolution of the boundary points. The two systems of equations are coupled and therefore must be used together in any technique attempting to find a solution for the continuous modelling problem M directly.

We will now show that in our discrete approximate model $M[D]$ the situation is different. While we still formally have a similar situation with internal and boundary points and different equations for each class of nodes, we will show that for our approximate model it is possible to derive a system of dynamic equations for all nodes.

Indeed, let us consider the following picture Figure 6.2 below of a typical solid (the solid is two dimensional for simplicity).

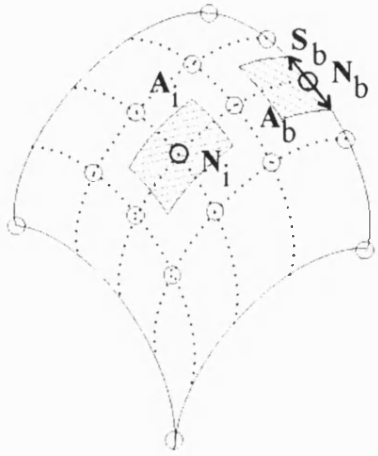


Figure 6.2: Solid with discrete nodes

In this picture, we have two nodes: an internal node $\mathbf{N}_i$ and a boundary node $\mathbf{N}_b$; $\mathbf{S}_b$ is the boundary area associated with the boundary node $\mathbf{N}_b$.

Since we are working with a discrete model and therefore have a finite number of nodes representing our body, each internal node has a finite piece of volume associated with it. We have shown this volume element as A_i.

Our dynamic equation describing the evolution of the node $\mathbf{N}_i$ is of the following form:

$$\ddot{\mathbf{R}}(\mathbf{N}_i,t)\rho(\mathbf{N}_i,t)=\mathbf{F}(\mathbf{N}_i,t)+\mathbf{E}(\mathbf{N}_i,t)$$

In our finite approximation, $\ddot{\mathbf{R}}(\mathbf{N}_i,t)$ refers to the acceleration of the centre of mass of this volume element A_i (we take it approximately to be the position of the node itself $\mathbf{R}(\mathbf{N}_i,t)$), $\rho(\mathbf{N}_i,t)$ is the average density of the volume element, so that the mass is given by:

$$m(\mathbf{N}_i,t)=\rho(\mathbf{N}_i,t)A_i \qquad (6.40)$$

The external volume force $\mathbf{F}(\mathbf{N}_i,t)$ and the elastic volume force $\mathbf{E}(\mathbf{N}_i,t)$ are actually averaged forces over the volume element; multiplied by the volume they will represent the actual forces acting on the volume element. This is really just a straightforward interpretation of the dynamic equations for the internal points of our discrete model.

We will now consider the boundary point $\mathbf{N}_b$. We can clearly see that, unlike in the continuous model, where boundary points do not have any volume associated with them, in our discrete model, we do have a finite volume element A_b associated with our boundary point. In the continuous model, this volume element A_b will always be associated with some internal point. In our finite discrete model, however, there is simply no internal node to associate with this volume element; it must therefore be assigned to the boundary node $\mathbf{N}_b$. Otherwise, there will be a whole layer of unaccounted volume along the boundary.

We can see that, in our discrete model, we can associate a volume element with every boundary node (if fact, we must do so if we are to cover the whole volume of the solid). If there is a volume element, it has density and therefore mass, and we can obtain dynamic equations for these boundary volume elements. As far as our discrete model is concerned, these dynamic equations will be formulated for the boundary points. The static conditions at

the boundary will be used implicitly in this formulation (more precisely, in the formula of the elastic force for the boundary points).

As the result of this, we will obtain a system of dynamic equations for all points in our discrete model, both internal and boundary. The static conditions will be used implicitly in the dynamic equations for the boundary points. We will therefore have a system of equations of the required form, without explicit static boundary conditions.

This system can then be converted into a first order differential equation in time and the discretization in time can be presented, together with the numerical techniques to solve the discretized system.

The existence of the boundary volume elements associated with boundary points and therefore dynamic equations for them is a feature unique to our discrete approximate model. In the continuous model there is no notion of boundary volume elements; instead the static boundary conditions must be considered explicitly to account for the external boundary surface forces.

6.1.3 Calculation of the elastic force for the boundary volume elements

We will now provide a complete description of the derivation of the expression of the elastic force for the boundary volume elements.

We will present the analysis in a two-dimensional case for simplicity.

Let us consider a boundary node $\mathbf{N}_b$ (see picture Figure 6.3). As we have just identified, there is a finite volume element A_b associated with the node $\mathbf{N}_b$.

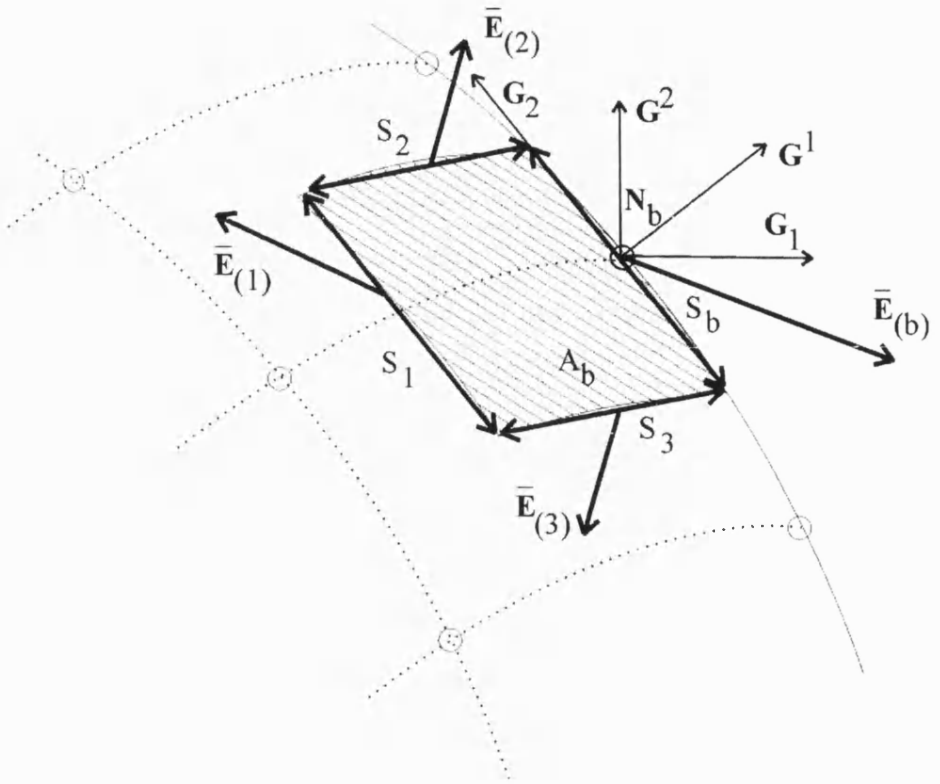


Figure 6.3: Elastic force derivation for a boundary volume element

The volume element is a parallelogram with the sides: S_b, S_1, S_2 and S_3, such that:

$$\begin{aligned} S_b &= S_1 \\ S_2 &= S_3 \end{aligned} \tag{6.41}$$

At the boundary point $\mathbf{N}_b$, we have the covariant basis $\mathbf{G}_i$ and the contravariant basis $\mathbf{G}^i$.

A force acts on each side of the volume element; their resultants for each node are marked as $\bar{\mathbf{E}}_{(b)}$, $\bar{\mathbf{E}}_{(1)}$, $\bar{\mathbf{E}}_{(2)}$ and $\bar{\mathbf{E}}_{(3)}$. $\bar{\mathbf{E}}_{(b)}$ is effectively the external boundary force and the other three forces $\bar{\mathbf{E}}_{(i)}$ are the internal reactionary forces, due to the body deformation. The sum $\bar{\mathbf{E}}$ of these four forces will be the resultant elastic force, acting in the boundary volume element:

$$\bar{\mathbf{E}} = \bar{\mathbf{E}}_{(b)} + \bar{\mathbf{E}}_{(1)} + \bar{\mathbf{E}}_{(2)} + \bar{\mathbf{E}}_{(3)} = \left(\bar{\mathbf{E}}_{(b)} + \bar{\mathbf{E}}_{(1)}\right) + \left(\bar{\mathbf{E}}_{(2)} + \bar{\mathbf{E}}_{(3)}\right) \tag{6.42}$$

Let us assume the current point is (l,m). We can now write down the expressions for all the forces involved. First of all the boundary force:

$$\overline{E}^i_{(b)}(l,m) = S_b S^i(l,m) \tag{6.43}$$

Here S_b is the surface area and $S^i(l,m)$ is the external boundary force.

All the other forces are:

$$\begin{aligned}
\overline{E}^i_{(1)}(l,m) &= -S_1\sigma^{i1}\left(l-\frac{1}{2},m\right) \\
\overline{E}^i_{(2)}(l,m) &= S_2\sigma^{i2}\left(l,m+\frac{1}{2}\right) \\
\overline{E}^i_{(3)}(l,m) &= -S_3\sigma^{i2}\left(l,m-\frac{1}{2}\right)
\end{aligned} \tag{6.44}$$

Expressions like $l-\frac{1}{2}$ are the equivalent of $x-\frac{1}{2}dx$ in the continuous case. Here we used the fact that the force transmitted via a surface element dA_j is $\sigma^{ij}dA_j\mathbf{G}_i$. In the case of $\overline{E}^i_{(1)}$, for example, $\mathbf{dA} = S_1\mathbf{G}^1$ and thus we obtain the expression for $\overline{E}^i_{(1)}$.

In the case of $\overline{E}^i_{(1)}(l,m)$, we use the fact that the stress tensor on the boundary is defined as the continuous extension of the stress at the internal nearby points. Therefore:

$$\overline{E}^i_{(1)}(l,m) = -S_1\sigma^{i1}\left(l-\frac{1}{2},m\right) = -S_1\sigma^{i1}(l,m) \tag{6.45}$$

For each of the other forces we can use the first two terms of the Taylor expansion to obtain:

$$\begin{aligned}
\overline{E}^i_{(2)}(l,m) &= S_2\sigma^{i2}(l,m) + S_2\sigma^{i2}\big|_2(l,m)\frac{1}{2}\Delta m \\
\overline{E}^i_{(3)}(l,m) &= -S_3\sigma^{i2}(l,m) + S_3\sigma^{i2}\big|_2(l,m)\frac{1}{2}\Delta m
\end{aligned} \tag{6.46}$$

If we now consider the sums $\left(\overline{\mathbf{E}}_{(b)} + \overline{\mathbf{E}}_{(1)}\right)$ and $\left(\overline{\mathbf{E}}_{(2)} + \overline{\mathbf{E}}_{(3)}\right)$ separately, we obtain the following equalities:

$$\left(\overline{\mathbf{E}}_{(b)} + \overline{\mathbf{E}}_{(1)}\right) = S_b S^i(l,m) - S_1\sigma^{i1}(l,m) = S_b\left(S^i(l,m) - \sigma^{i1}(l,m)\right) \tag{6.47}$$

The surface area S_b is connected to the volume of the volume element via the following approximate expression:

$$S_b \cong \frac{A_b}{\Delta l/2} = \frac{2A_b}{\Delta l} \tag{6.48}$$

Therefore:

$$\left(\overline{\mathbf{E}}_{(b)} + \overline{\mathbf{E}}_{(1)}\right) \cong \frac{2A_b}{\Delta l}\left(S^i(l,m) - \sigma^{i1}(l,m)\right) \tag{6.49}$$

And now the other sum:

$$\begin{aligned}\left(\overline{\mathbf{E}}_{(2)} + \overline{\mathbf{E}}_{(3)}\right) &= S_2\sigma^{i2}(l,m) + S_2\sigma^{i2}\big|_2(l,m)\frac{1}{2}\Delta m - \\ &- S_3\sigma^{i2}(l,m) + S_3\sigma^{i2}\big|_2(l,m)\frac{1}{2}\Delta m = S_2\sigma^{i2}\big|_2(l,m)\Delta m\end{aligned} \tag{6.50}$$

The surface area of the volume element S_2 is connected to the volume of the volume element as follows:

$$S_2 \cong \frac{A_b}{\Delta m} \tag{6.51}$$

Therefore:

$$\left(\overline{\mathbf{E}}_{(2)} + \overline{\mathbf{E}}_{(3)}\right) \cong A_b\sigma^{i2}\big|_2(l,m) \tag{6.52}$$

The resultant elastic force $\overline{\mathbf{E}}$ for the whole boundary volume element is finally:

$$\begin{aligned}\overline{E}^i &\cong \frac{2A_b}{\Delta l}\left(S^i(l,m) - \sigma^{i1}(l,m)\right) + A_b\sigma^{i2}\big|_2(l,m) = \\ &= A_b\left(\frac{2}{\Delta l}\left(S^i(l,m) - \sigma^{i1}(l,m)\right) + \sigma^{i2}\big|_2(l,m)\right)\end{aligned} \tag{6.53}$$

Similar expressions can be obtained for all other possible boundary situations: corners, edges and side planes (in the full three-dimensional case).

We must note that all of the over-barred forces are the full forces: $\overline{\mathbf{E}}_{(b)}$ and $\overline{\mathbf{E}}_{(i)}$ are the forces acting on the corresponding surface elements, and $\overline{\mathbf{E}}$ is the force acting on the whole

volume of the volume element. To convert the volume elastic force into the averaged one, that will take part in the equations of motion, we need to divide it over the volume:

$$E^i = \frac{\overline{E}^i}{A_b} = \frac{2}{\Delta l}\left(S^i(l,m) - \sigma^{i1}(l,m)\right) + \sigma^{i2}\Big|_2(l,m) \tag{6.54}$$

We have now shown the procedure for obtaining the expressions of the elastic forces for boundary volume elements (which we have associated with the boundary nodes). Combining these with the expressions for the purely internal nodes, we, finally, arrived to the situation where no explicit static boundary conditions are needed; we have used them implicitly in the expressions for the elastic forces of the boundary points.

Having derived the expression for the elastic force of the boundary volume element, we can combine it with the inertial force and an external volume force, acting on the volume element, to obtain the full system of dynamic equations for the boundary volume element:

$$\ddot{\mathbf{R}}(l,m,n,t)\rho(l,m,n,t) = \mathbf{F}(l,m,n,t) + \mathbf{E}(l,m,n,t) \tag{6.55}$$

The system of equations looks identical to the ones used in the main system of equations (6.24), with only the term $\mathbf{E}(l,m,n,t)$ for the elastic force, having a different expression.

The full system of equations governing the evolution of the solid is now of the following form:

$$\begin{cases} \ddot{\mathbf{R}}(l,m,n,t)\rho(l,m,n,t) = \mathbf{F}(l,m,n,t) + \mathbf{E}(l,m,n,t) \\ \mathbf{R}(l,m,n,0) = \mathbf{R}(l,m,n) \\ \dot{\mathbf{R}}(l,m,n,0) = \mathbf{V}(l,m,n) \\ 0 \le l \le L, 0 \le m \le M, 0 \le n \le N \end{cases} \tag{6.56}$$

The elastic force $\mathbf{E}(l,m,n,t)$ is calculated differently depending on whether it is an internal node or a boundary one.

6.2 Numerical solution of the discrete system of equations

As the first step, we need to convert the system of the ordinary differential equations of the second order:

$$\begin{cases} \ddot{\mathbf{R}}(l,m,n,t)\rho(l,m,n,t) = \mathbf{F}(l,m,n,t) + \mathbf{E}(l,m,n,t) \\ \mathbf{R}(l,m,n,0) = \mathbf{R}(l,m,n) \\ \dot{\mathbf{R}}(l,m,n,0) = \mathbf{V}(l,m,n) \\ 0 \le l \le L, 0 \le m \le M, 0 \le n \le N \end{cases} \tag{6.57}$$

into a system of the ordinary differential equations of the first order in the following canonical form:

$$\begin{cases} \dot{\mathbf{r}}(t) = \mathbf{f}(t, \mathbf{r}(t)), t > 0 \\ \mathbf{r}(0) = \mathbf{r}_0 \end{cases} \tag{6.58}$$

We do it using the usual procedure. We introduce a new vector variable:

$$\begin{aligned} \mathbf{\Lambda}(l,m,n,t) &= \begin{pmatrix} \mathbf{V}(l,m,n,t) \\ \mathbf{R}(l,m,n,t) \end{pmatrix} \\ \mathbf{\Lambda}(l,m,n) &= \begin{pmatrix} \mathbf{V}(l,m,n) \\ \mathbf{R}(l,m,n) \end{pmatrix} \end{aligned} \tag{6.59}$$

And a new vector function (corresponding to the right part $\mathbf{f}(t, \mathbf{r}(t))$ of the canonical form (6.58)):

$$\mathbf{\Omega}(l,m,n,t) = \begin{pmatrix} \dfrac{1}{\rho(l,m,n,t)} (\mathbf{F}(l,m,n,t) + \mathbf{E}(l,m,n,t)) \\ \mathbf{V}(l,m,n,t) \end{pmatrix} \tag{6.60}$$

Using this new notation and the new variable, the system will take the following form:

$$\begin{cases} \dot{\mathbf{\Lambda}}(l,m,n,t) = \mathbf{\Omega}(l,m,n,t), t > 0 \\ \mathbf{\Lambda}(l,m,n,0) = \mathbf{\Lambda}(l,m,n) \\ 0 \le l \le L, 0 \le m \le M, 0 \le n \le N \end{cases} \tag{6.61}$$

This system of equation is the system of ordinary differential equations of the first order with an initial condition, with regard to the unknown vector function $\mathbf{\Lambda}(l,m,n,t)$.

Before we proceed to present the numerical solution scheme for this system of ordinary equations, it is useful to make the following note. The right part $\mathbf{\Omega}(l,m,n,t)$ is really a function of all positions of the solid nodes at the time instant t, rather than the indices

(l,m,n). Indeed, the external force and the density can be formally presented as function of the nodes:

$$\begin{aligned}\mathbf{F}(l,m,n,t)&=\mathbf{F}(\mathbf{\Lambda}(l,m,n,t))\\ \rho(l,m,n,t)&=\rho(\mathbf{\Lambda}(l,m,n,t))\end{aligned} \tag{6.62}$$

The elastic force, however, is a function of all nodes of the solid, not just the current one (l,m,n,t), therefore:

$$\mathbf{\Omega}(l,m,n,t)=\mathbf{\Omega}(\mathbf{\Lambda}(0,0,0,t),...,\mathbf{\Lambda}(o,p,q,t),...,\mathbf{\Lambda}(L,M,N,t)) \tag{6.63}$$

Or, for brevity:

$$\mathbf{\Omega}(l,m,n,t)=\mathbf{\Omega}(\mathbf{\Lambda}(o,p,q,t)) \tag{6.64}$$

Where, due to the fact that a different set of indices is used on the right, $\mathbf{\Omega}(\mathbf{\Lambda}(o,p,q,t))$ is shorthand for the full version $\mathbf{\Omega}(\mathbf{\Lambda}(0,0,0,t),...,\mathbf{\Lambda}(o,p,q,t),...,\mathbf{\Lambda}(L,M,N,t))$ but with the same meaning.

Now the system of the equations is formally transformed into the following equation (6.65):

$$\begin{cases}\dot{\mathbf{\Lambda}}(l,m,n,t)=\mathbf{\Omega}(\mathbf{\Lambda}(o,p,q,t)),t>0\\ \mathbf{\Lambda}(l,m,n,0)=\mathbf{\Lambda}(l,m,n)\\ 0\le l\le L,0\le m\le M,0\le n\le N\end{cases} \tag{6.65}$$

We can now proceed with the actual discretization in time.

As usual, instead of having a continuous time $t>0$, we consider a sequence $\{t_k\}$, such as:

$$t_k\ge 0,t_0=0,t_k<t_{k+1} \tag{6.66}$$

Any numerical scheme has to answer the following two questions:

- How to choose the time step when moving from the current time instant t_k to the next one t_{k+1}, that is, how to choose $t_{k+1}-t_k$.
- How to advance the whole system to the next time instant t_{k+1}.

Both answers are specific to a particular numerical scheme. But whatever the numerical system, it is clear what the general procedure is as follows: we have the initial condition $\Lambda(l,m,n,0)=\Lambda(l,m,n)$, which is effectively the state of the system at the initial time instant $t_0=0$. We can now determine what the next step t_1 will be, and then move the whole system to this time instant. We can then apply the same procedure again and again, to move the system forward in time from one time instant to another.

6.2.1 Numerical scheme overview

Generally, all the numerical schemes are broadly divided into a number of classes:

- Explicit and implicit methods. In explicit numerical schemes, the system is advanced to the next step directly, using only the information from the previous step(s) (and possibly at some intermediate points as well). In implicit numerical schemes, the information about the state of the system at the next step is contained in some algebraic formulae. Therefore, to find the state of the system at the next step, a system of (normally linear and algebraic) equations must be solved.

- Methods with time step control and those without. The schemes that provide means to set variable time step (as opposed to a constant time interval) are normally more advanced. The step size is normally chosen to reflect the system behaviour in some way to prevent instability and a severe loss of precision.

- Multi-step or single-step methods. In single-step numerical schemes, only the information from the previous step is used in the calculation of the state of the system at the next step; while in multi-step schemes, not only the previous time step but also other system states in the past are used to evaluate the state of the system at the next time step.

- Higher order methods. In these methods more intermediate evaluations are used to obtain the system state at the nest time step. These methods are normally more precise but also more computationally expensive.

- Hybrid methods. These methods combine the features from more than one class to create a method that is more adapted to particular applications.

Each of these classes is considered separately in Appendix 1.

In this thesis we have used the 5th order Runge-Kutta method with step control. We found it to be well suited for our modelling problem: it is fast, stable, adequately accurate and relatively easy to implement. In the next section we present this method in the form that was used in this thesis for all practical implementations

6.2.2 The 5th order Runge-Kutta method with step control

The 5th order Runge-Kutta method with step control is an evolved technique for solving the first order systems of equations. It is quite accurate, has good support for error evaluation and time step adjustment, and is quite easy to implement and use. See [100] for its complete and detailed description.

The method uses the following scheme to advance the solution to the next step:

$$\begin{aligned}
\mathbf{k}_1(l,m,n,t_k) &= \Delta t_k[\boldsymbol{\Omega}(\boldsymbol{\Lambda}(o,p,q,t_k))] \\
\mathbf{k}_2(l,m,n,t_k) &= \Delta t_k[\boldsymbol{\Omega}(\boldsymbol{\Lambda}(o,p,q,t_k)+b_{21}\mathbf{k}_1(o,p,q,t_k))] \\
\mathbf{k}_3(l,m,n,t_k) &= \Delta t_k\left[\boldsymbol{\Omega}\left(\boldsymbol{\Lambda}(o,p,q,t_k)+\sum_{i=1}^{2} b_{3i}\mathbf{k}_i(o,p,q,t_k)\right)\right] \\
\mathbf{k}_4(l,m,n,t_k) &= \Delta t_k\left[\boldsymbol{\Omega}\left(\boldsymbol{\Lambda}(o,p,q,t_k)+\sum_{i=1}^{3} b_{4i}\mathbf{k}_i(o,p,q,t_k)\right)\right] \\
\mathbf{k}_5(l,m,n,t_k) &= \Delta t_k\left[\boldsymbol{\Omega}\left(\boldsymbol{\Lambda}(o,p,q,t_k)+\sum_{i=1}^{4} b_{5i}\mathbf{k}_i(o,p,q,t_k)\right)\right] \\
\mathbf{k}_6(l,m,n,t_k) &= \Delta t_k\left[\boldsymbol{\Omega}\left(\boldsymbol{\Lambda}(o,p,q,t_k)+\sum_{i=1}^{5} b_{6i}\mathbf{k}_i(o,p,q,t_k)\right)\right] \\
\boldsymbol{\Lambda}(l,m,n,t_{k+1}) &= \boldsymbol{\Lambda}(l,m,n,t_k)+\sum_{i=1}^{6} c_i\mathbf{k}_i(l,m,n,t_k)
\end{aligned} \tag{6.67}$$

This method does evaluations at six different intermediate points and then takes their weighted average to approximate the state of the system at the next step.

The error estimate is given by the following formula:

$$\Delta\boldsymbol{\varepsilon}_k = \sum_{i=1}^{6}(c_i - \bar{c}_i)\mathbf{k}_i \tag{6.68}$$

The error is of $O\left(|\Delta t_k|^5\right)$ order with the regard to the step size:

$$|\Delta\boldsymbol{\varepsilon}_k| = O\left(|\Delta t_k|^5\right) \tag{6.69}$$

Constants b_{ij} are given in the following table:

b_{ij}	1	2	3	4	5
1					
2	$\frac{1}{5}$				
3	$\frac{3}{40}$	$\frac{9}{40}$			
4	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$		
5	$-\frac{11}{54}$	$\frac{5}{2}$	$-\frac{70}{27}$	$\frac{35}{27}$	
6	$\frac{1631}{55296}$	$\frac{175}{512}$	$\frac{575}{13824}$	$\frac{44275}{110592}$	$\frac{253}{4096}$

(6.70)

Constants c_i and $\bar{c}_i$ are given in the following table:

i	c_i	$\bar{c}_i$
1	$\frac{37}{378}$	$\frac{2825}{27648}$
2	0	0
3	$\frac{250}{621}$	$\frac{18575}{48384}$

(6.71)

4	$\frac{125}{594}$	$\frac{13525}{55296}$
5	0	$\frac{277}{14336}$
6	$\frac{512}{1771}$	$\frac{1}{4}$

Having determined the error estimate $\Delta\boldsymbol{\varepsilon}_k$, the corrected step to keep the error at the level of the preset error threshold A is given by the following equality:

$$\Delta\bar{t}_k = \Delta t_k \left| \frac{\mathrm{A}}{\|\Delta\varepsilon\|} \right|^{\frac{1}{5}} \tag{6.72}$$

Where $|\Delta\varepsilon|$ is the scalar metric of the vector error estimates $\Delta\boldsymbol{\varepsilon}_k$.

The adjusted step size $\Delta\bar{t}_k$ is the one that is required to keep the error under control (at the level of the desired accuracy A).

6.3 Practical implementation of the numerical solution

6.3.1 Adjusting the time step

Theoretically, having calculated the new step size, it should be used to recalculate $\boldsymbol{\Lambda}(o,p,q,t_{k+1})$. In practice, however, we found that it is quite expensive computationally and instead of recalculating $\boldsymbol{\Lambda}(o,p,q,t_{k+1})$ with the new time step, we update the time step (for the next round) but do not recalculate $\boldsymbol{\Lambda}(o,p,q,t_{k+1})$.

In this scenario, the new (adjusted) time step will be used during the evaluation of the next system state $\boldsymbol{\Lambda}(o,p,q,t_{k+2})$. The time step is adjusted with a delay and not immediately. Theoretically, this may lead to instability in certain extreme cases, but we did not find it to be a problem in practice.

In fact, we found the numerical solution with the time step size adjusted in this way to be extremely robust and stable. It was able to cope with extreme deformations and perturbations.

6.3.2 Choosing the metric for the error estimate

The error estimation is a vector $\Delta\boldsymbol{\varepsilon}_k = \Delta\boldsymbol{\varepsilon}(l,m,n,t_k)$. In the step correction formula we need to calculate its metric $|\Delta\boldsymbol{\varepsilon}_k|$. Theoretically, any metric can be used to obtain this scalar, for example, the vector length:

$$|\Delta\boldsymbol{\varepsilon}_k| = \sqrt{\sum_i \left(\Delta(\varepsilon^i)_k\right)^2} \tag{6.73}$$

The notation $\Delta(\varepsilon^i)_k$ denotes the contravariant components of the vector $\Delta\boldsymbol{\varepsilon}_k$:

$$\Delta\boldsymbol{\varepsilon}_k = \sum_i \Delta(\varepsilon^i)_k \mathbf{k}_i \tag{6.74}$$

Where $\mathbf{k}_i$ are the Cartesian basis of our global coordinate system.

However, in practice we found it is easier and less computationally expensive to use another metric to obtain a scalar measure of the error:

$$|\Delta\boldsymbol{\varepsilon}_k| = \max_i |\Delta(\varepsilon^i)_k| \tag{6.75}$$

Having found the metric for $|\Delta\boldsymbol{\varepsilon}_k| = |\Delta\boldsymbol{\varepsilon}(l,m,n,t_k)|$, the overall error can be calculated as a further maximum of the scalar error estimate at all nodes:

$$\Delta\varepsilon = \max_{l,m,n} |\Delta\boldsymbol{\varepsilon}(l,m,n,t_k)| \tag{6.76}$$

This max metric not only can be used to adjust the step, but also to provide a visual estimate of the evaluation error in a simple scalar form at any step.

6.3.3 Dynamic incremental alignment of the reference frame

This is one topic we have not explained in detail before and will do it now.

At every time step t_k and for every point (l,m,n) we need the reference frame and the deforming body to be aligned both translationally and rotationally. The deforming body has been denoted as $\mathbf{R}(l,m,n,t_k)$ and the aligned reference frame as $\mathbf{R}'(l,m,n)$.

Translational alignment simply means we matched both bodies at the alignment point:

$$\mathbf{R}(l,m,n,t_k) = \mathbf{R}'(l,m,n) \tag{6.77}$$

We know that the criteria for the correct rotational alignment is expressed by the following equation:

$$curl(\mathbf{U}(l,m,n,t_k)) = 0 \tag{6.78}$$

Where the displacement field $\mathbf{U}(l,m,n,t_k)$ is calculated as usual:

$$\mathbf{U}(l,m,n,t_k) = \mathbf{R}(l,m,n,t_k) - \mathbf{R}'(l,m,n) \tag{6.79}$$

The fact that the curl of the displacement field is zero at the centre of alignment (point (l,m,n)) means that the displacement field does not contain a rotational component and is therefore purely deformational.

When the curl of the displacement field is not zero, it means the displacement field does have a rotational component:

$$\mathbf{\Theta} = \frac{1}{2} curl(\mathbf{U}(l,m,n,t_k)) \tag{6.80}$$

The vector $\mathbf{\Theta}$ will be its rotational vector: its length will be the actual angle of rotation and its direction will be perpendicular to the plane of rotation. We will now show that this equality is true for any angles $\mathbf{\Theta}$, not only for small.

Indeed

$$curl(\mathbf{U} + \mathbf{w} \times \mathbf{r}) = curl\mathbf{U} + curl(\mathbf{w} \times \mathbf{r}) = curl\mathbf{U} + 2\mathbf{w} \tag{6.81}$$

Where $\mathbf{r}$ is the position vector and $\mathbf{w}$ is any rotational vector.

We will now present the actual dynamic alignment procedure.

At the beginning of the modelling experiment at $t = t_0 = 0$ we have (from the initial condition):

$$\mathbf{R}(l,m,n,0) = \mathbf{R}(l,m,n,t_0) = \mathbf{R}(l,m,n) \tag{6.82}$$

This means that at $t = t_0 = 0$ the solid and its reference frame are guaranteed to be perfectly aligned.

We proceed to the next step t_1. We have the deforming solid in its new position $\mathbf{R}(l,m,n,t_1)$, the reference frame is still where it was at the start.

The alignment procedure for the step t_1 will be as follows:

- First, we align the deforming solid and the reference frame transitionally, we will obtain $\overline{\mathbf{R}}(l,m,n)$: $\mathbf{R}(l,m,n,t_1)=\overline{\mathbf{R}}(l,m,n)$

- We then calculate the displacement field away from $\overline{\mathbf{R}}(l,m,n)$:

$$\overline{\mathbf{U}}(l,m,n,t_1)=\mathbf{R}(l,m,n,t_1)-\overline{\mathbf{R}}(l,m,n) \tag{6.83}$$

  And then its curl at every point:

$$\mathbf{\Theta}=\frac{1}{2}curl\left(\overline{\mathbf{U}}(l,m,n,t_1)\right) \tag{6.84}$$

  Because the solid and its reference frame $\overline{\mathbf{R}}(l,m,n)$ are not aligned rotationally, $\mathbf{\Theta}$ is generally not zero.

  The rotational vector $\mathbf{\Theta}$ can then be used to rotate the reference frame $\overline{\mathbf{R}}(l,m,n)$ at every point (more precisely its local neighbourhood for every point) to obtain rotationally aligned $\mathbf{R}'(l,m,n)$.

  The correct displacement field (without the rotational component) can then be obtained as usual:

$$\mathbf{U}(l,m,n,t_1)=\mathbf{R}(l,m,n,t_1)-\mathbf{R}'(l,m,n) \tag{6.85}$$

This process can then be repeated for all other time steps.

6.3.4 Dissipation of energy (damping force)

If in real life we apply a constant volume force (like gravity) to a solid and have it fall onto a constraint (like floor), due to its elastic properties, the solid will bounce. This is due to the energy transformations: the kinetic energy transforms into the potential energy on impact, then this potential energy back to the kinetic energy on bounce and so on.

We do know, however, that any real solid will eventually stop bouncing. This physically will be due to various internal and external factors (i.e. heat), which could be summarised simply by saying that the solid's energy is dissipated during its deformation.

In our model, nothing yet accounts for any energy loss. The only possible source of the energy variations is in the rounding errors of the numerical solution: due to the finite precision of the numerical solution and rounding errors during the calculations, all the quantities are calculated with a certain small error, which in its turn can lead to minor fluctuations in the overall solid's energy balance. These errors, however, are kept under control by the step adjustments; if they were not, this would quickly lead to instability.

To account for the energy loss due to external factors, we introduce a notion of the dissipation force $\mathbf{D}$. This is a volume force and is most closely resembles the physical internal friction force.

We first of all define the dissipation force acting through any area element inside the solid:

$$D_i = -D\dot{\varepsilon}_{ij} dA^j \tag{6.86}$$

Where D is the dissipation constant, if it is zero there is no dissipation force, and dA^j is the area element, resolved into contravariant components.

Using the same arguments as with the derivation of the expression of the elastic force for internal volume elements, we can write that for any internal volume element, the dissipation force acting on it will be:

$$\begin{aligned} D_i &= -D\dot{\varepsilon}_{ij}\Big|^j \Leftrightarrow \\ D^i &= -D\dot{\varepsilon}^{ij}\Big|_j \end{aligned} \tag{6.87}$$

As with any other volume force, the dissipation force must be added to the resultant $\mathbf{F}(l,m,n,t_k)$ (after a suitable approximation has been obtained for this continuous case).

6.3.5 Additional surface forces: external surface friction and wind

There are two types of external forces that can act on the solid at any time.

The first type consists of the external volume forces, which for most practical applications include only gravity. These are represented by their resultant $\mathbf{F}(l,m,n,t_k)$ and are expressed

in force per unit volume. If we wish to add an external force of this kind, we simply need to average it per unit volume and add it to the resultant $\mathbf{F}(l,m,n,t_k)$: the numerical solution will take it into account. We can even do it after the animation has started (for example let the user turn on/off the gravity), and as far as the numerical solution is concerned, it simply needs a vector $\mathbf{F}(l,m,n,t_k)$ specified at all times, regardless of how it has been obtained.

The second type consists of the surface forces acting through the solid's boundary, and these include all contact forces, like surface friction and wind (air resistance). These forces are specified by their resultant $\mathbf{S}(l,m,n,t_k)$ but, unlike the volume forces, the surface forces are applied to the boundary area, which is represented by a vector perpendicular to the area's surface and are therefore dependant on the direction of the surface area element as well as the size of the area they act upon. As in the case of the volume forces, the surface forces are expressed in force units per unit area.

The volume forces are relatively straightforward to add and manipulate. We will therefore turn our attention to the boundary forces. We will show how boundary forces can be added on the example of two different surface forces: the friction force and the force produced by the wind due to the air resistance.

First we will derive the expression for the friction force. We use a simple fact that for slow moving bodies, the force of friction between the areas of our solid and the external object that are in direct contact is given by the following expression:

$$\mathbf{S}_{fr} = -\alpha \cdot \mathbf{v}_p \qquad (6.88)$$

Where $\mathbf{v}_p$ is the velocity of our solid relative to the external object and α is the friction coefficient (see Figure 6.4 below).

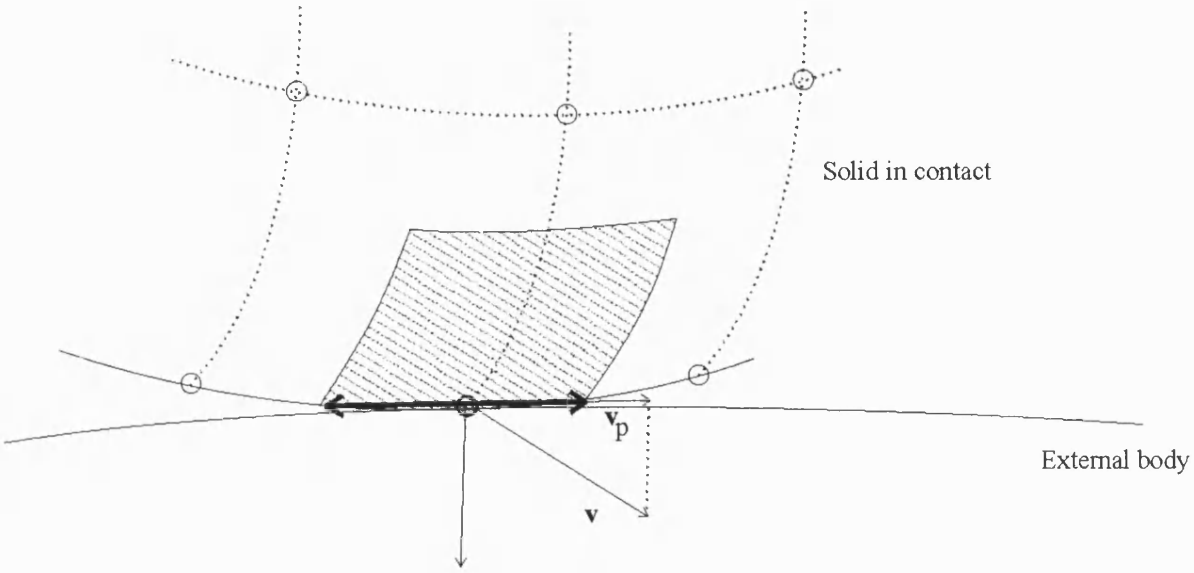


Figure 6.4: Friction force

Practically, we need to obtain the normal vector to the area of contact, calculate the projection $\mathbf{v}_p$ of the velocity $\mathbf{v}$ of the point in contact and then use the formula to calculate the friction force $\mathbf{S}_{fr}$. It can then be added to the resultant $\mathbf{S}(l,m,n,t_k)$ and the numerical solution will take it into account.

We will now derive the expression for the force of wind.

Let us assume we have wind of constant strength, represented by a vector $\mathbf{W}$. The body is submerged into the stream (see Figure 6.5 below).

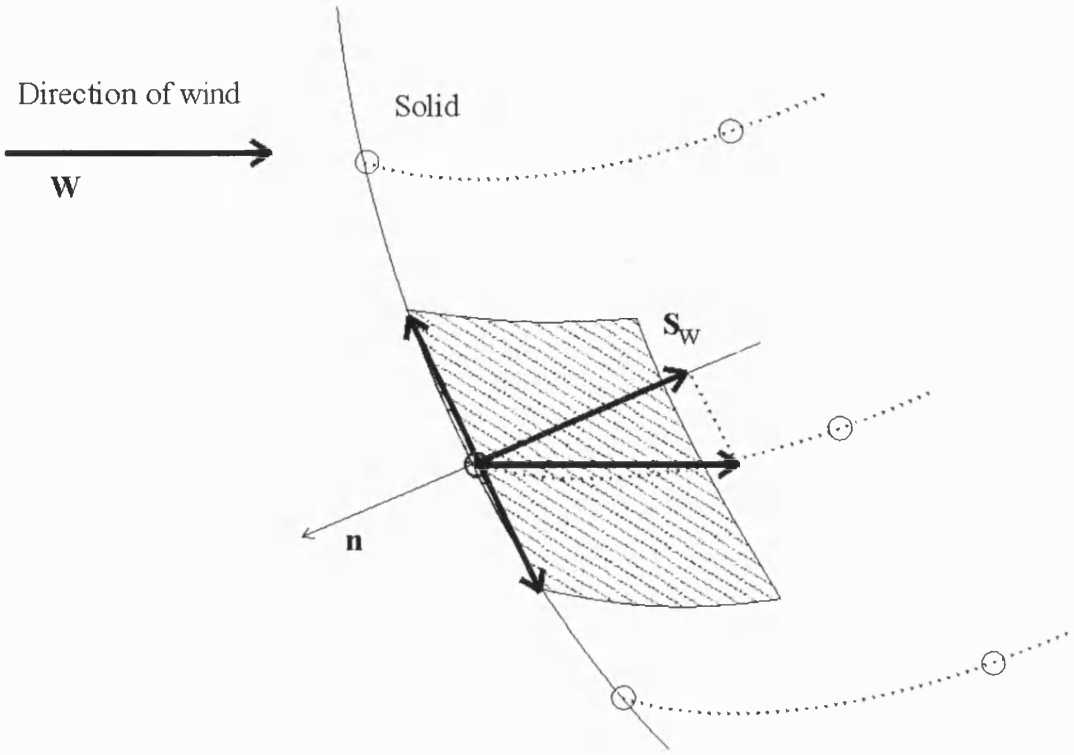


Figure 6.5: Wind

The force of wind is defined as the projection of the strength of the wind onto the internal normal (opposite to the external normal) to the solid's boundary:

$$\mathbf{S}_w = -(\mathbf{n} \cdot \mathbf{W})\mathbf{n} \qquad (6.89)$$

We can see from the definition of the wind force, it will be zero at the side of the solid that is not exposed directly to the wind. It will also be zero at the sides that are parallel to the wind. Both of these facts make physical sense for simple approximations of winds (when we do not take into account such things as turbulence).

As can be seen from the derivation of the expression of the wind force, it is not required to be constant as we assumed before. It can vary, and the same expression will work as long as we know (or can calculate) the direction vector $\mathbf{W}$ at every boundary point.

6.4 Examples

In this section we will present a range of examples to show our theory of FlexyMatter in action and prove its validity as a practical modelling technique.

For simplicity in all the examples below purely two-dimensional objects are used. Unless stated otherwise, there is a gravitational field directed downwards as usual. The black line at the bottom of some of the pictures is the floor – a rigid surface, which the animated solids

are prohibited from passing through. In all the examples we used a small non-zero value for the Poisson ratio ν.

Example 1: a solid suspended at an internal node

In this example, we are going to subject a square elastic solid to a severe deformation by suspending it at one of its nodes and gradually reducing the modulus of elasticity E. As E decreases, so do the internal forces that are responsible for maintaining the solid's geometric shape. Thus, large deformation should result for sufficiently low values of E.

As can be seen in Figure 6.6, the solid is suspended in the gravitational field at an internal node. On each step we halved the modulus of elasticity E and each step represents the equilibrium state of the system for the given value of E.

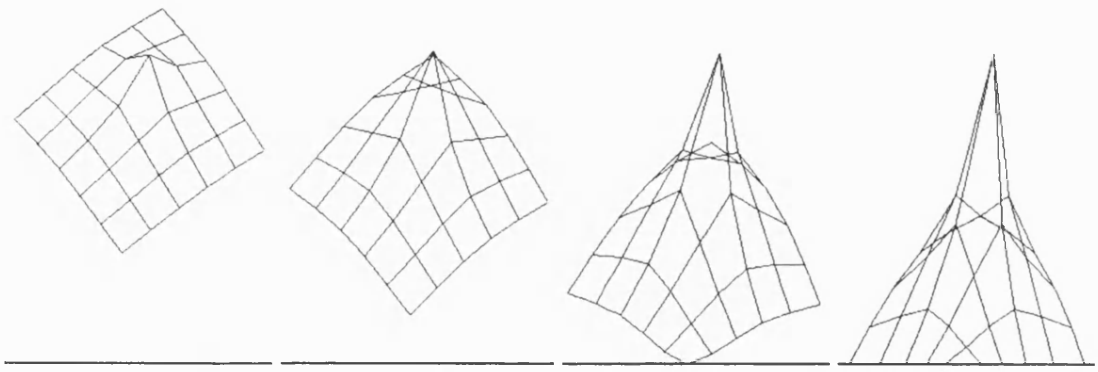


Figure 6.6: A square solid suspended at an internal node with the decreasing modulus of elasticity E (halved at each step)

It is interesting to see how our model copes with the situation when the neighbourhood of the suspended node gets more and more distorted until the mesh gets intermingled. It is particularly pronounced in the last picture of the sequence, where E is one sixteenth of the original value used to produce the first picture. In fact there is an illusion that the object becomes three-dimensional (a sheet of rubber suspended at a point). It is a pure illusion; the system remains purely two-dimensional, with a very intermingled mesh of nodes.

Example 2: a square solid stretched by the two opposite corner nodes

Now we are going to subject a solid to another constrained deformation. In such a way, we can obtain the equilibrium shapes of the solid under various levels of stress and compare it with our expectations from the real world.

We have a square solid in the gravitational field. We subject it to the following controlled deformation: we constrain two opposite corners and start to move them away from each other in opposite directions (see the sequence in Figure 6.7 below).

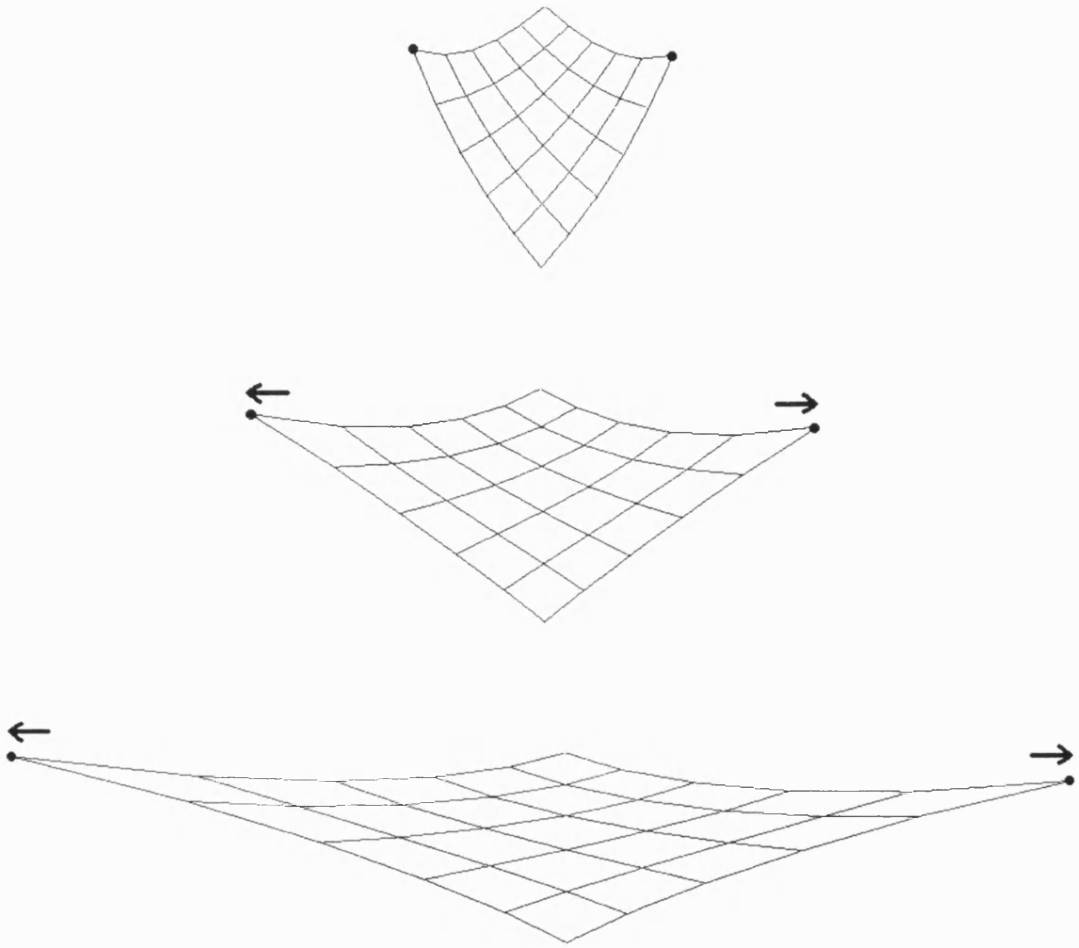


Figure 6.7: A solid suspended at the opposite corner nodes is stretched in opposite directions

We can see that the downward bulge becomes less pronounced as the solid gets stretched and the internal forces inside the solid grow. Finally, as the solid is stretched even more, the gravitational force becomes much smaller than the internal stresses inside the solid and therefore has little effect on the solid (we can see the stretched solid starts to assume a symmetric shape).

This effect matches well with our real-life experience.

Example 3: a solid stretched at two of its bottom corner nodes

This example is somewhat similar to the previous one in that is also has two constrained corner nodes. But this time the constrained nodes are the two adjacent bottom nodes of the

solid. In this configuration, the body of the solid is no longer evenly distributed above and under the line connecting the two nodes. Instead, we have the whole of the solid above the line pressing with all of its weight on the two constrained nodes.

Our intention is to see what will happen to the shape of the solid if we pull the nodes apart in opposite directions parallel to the floor. The following Figure 6.8 shows the sequence.

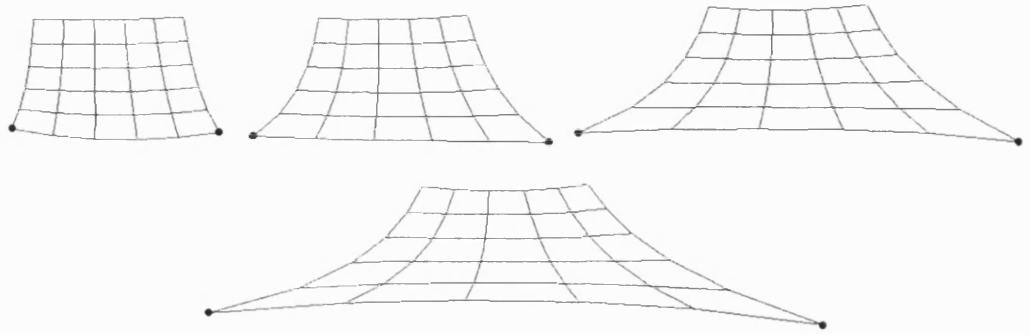


Figure 6.8: A solid stretched at two of its bottom corner nodes

As we move the two bottom nodes away from each other, we can see the bottom line starts to straighten up. This is in line with our expectations (see the previous example, for instance). But as we continue to move the nodes, we can see that the bottom line not only straightened up completely, but also started to bend the opposite way! On the last picture in the sequence, we can see that the bottom line of the solid has fully reversed its initial downward bulging and is now bulged upward. This is something that is not necessarily easy to predict from our real life experience.

The next two examples are relatively simple. The aim here is to examine the performance of the modelling process at different values of the modulus of elasticity E including very large and especially very small. It is particularly interesting to see what happens when very small values of E are set – in this case the internal stresses are very small compared to the external applied force (gravity). As the result solid's nodes can easily come very close to each other, including complete collapse of the solid on itself, or on the other hand get separated by a large distance. Increasing the value of E should restore the solid back to its normal shape.

Example 4: a solid suspended at its top right corner node

In this example we have a square solid with one (top right) corner node held fixed. Each of the example pictures in the main sequence was obtained in the following way: we start the

simulation in the original position (the top picture in the sequence Figure 6.9 below) and wait for the solid to settle down for the given values of the parameters (in our case E is varying). After the equilibrium was reached the picture was recorded.

We started with a certain initial value E_0 of E and on each step we halved it: $E = E_0, \frac{1}{2}E_0, \frac{1}{4}E_0, \ldots$. Figure 6.9 presents the sequence of pictures we obtained.

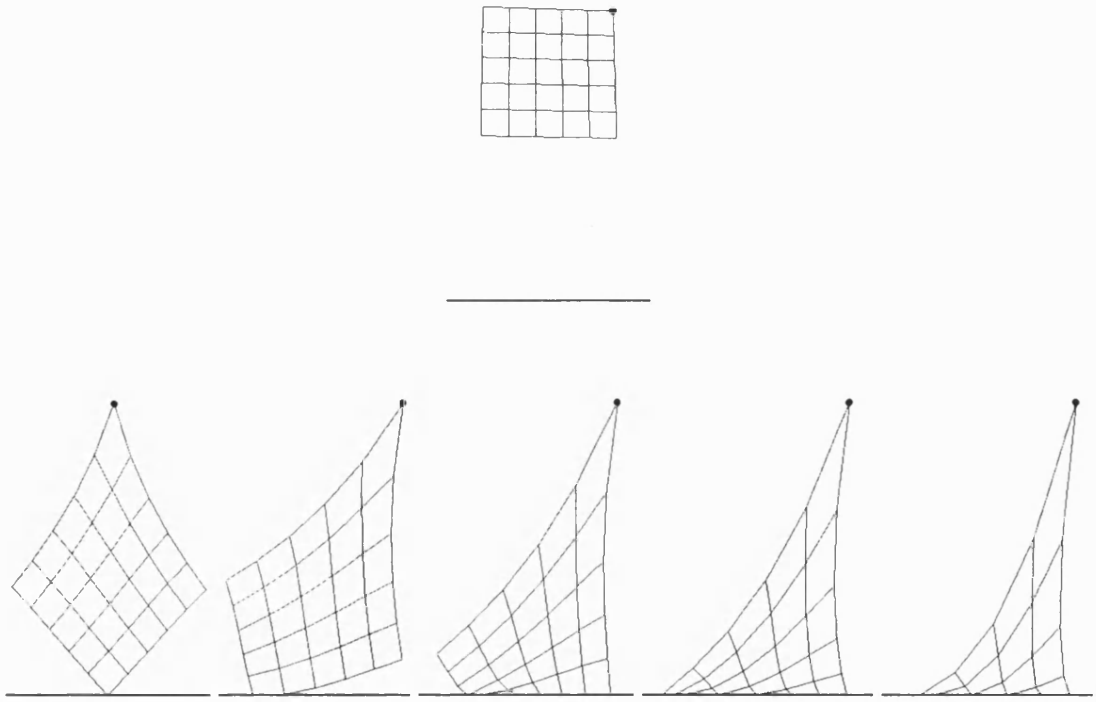


Figure 6.9: A solid suspended at its top right corner node with varying values of E (halved on each step)

We can clearly see the gradual collapse of the internal structure of the solid. As E is decreased, the internal stresses that take part in the internal elastic (restorative) force get weaker and less able to maintain the solid's internal geometrical structure.

This effect is even more pronounced in the following example.

Example 5: a free solid on the floor

In the next example we will imagine that we have a temperature dependent elastic substance that has the ability to increase its modulus of elasticity E as its temperature decreases.

We begin with the substance at a high temperature when $E \to 0$ (Figure 6.10).

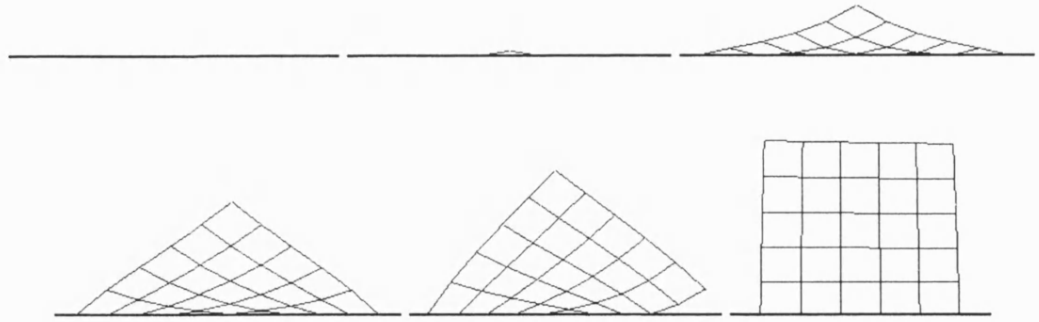


Figure 6.10: A temperature dependent substance. As it cools down, E increases.

The first picture shows the substance as a finite line element, held to its length by the effect of surface resistance. As the temperature decreases and E starts to grow ($E > 0$), we illustrate in the sequence the dynamic recovery of the substance.

This example demonstrates the versatility of our modelling technique. It is able to handle stably a wide range of situations, both constrained and unconstrained with a wide range of values for its internal parameters.

The following group of examples will present animation sequences demonstrating the deformations as they occurred during the simulations. These examples are different from the ones we presented above in that above every picture of every sequence represented the equilibrium state of the solid once the parameters were set and the animation started. The examples below will give raw animation sequences with the snapshots of the solid's evolution along its path towards an equilibrium state.

Example 6: a solid suspended at two of its top corner nodes is released and it collides with the floor

In this example our aim is to illustrate the dynamics of a solid, suspended at two of its top corner nodes and then released. A moderate value of the modulus of elasticity was used in this experiment to correspond to a fairly elastic solid.

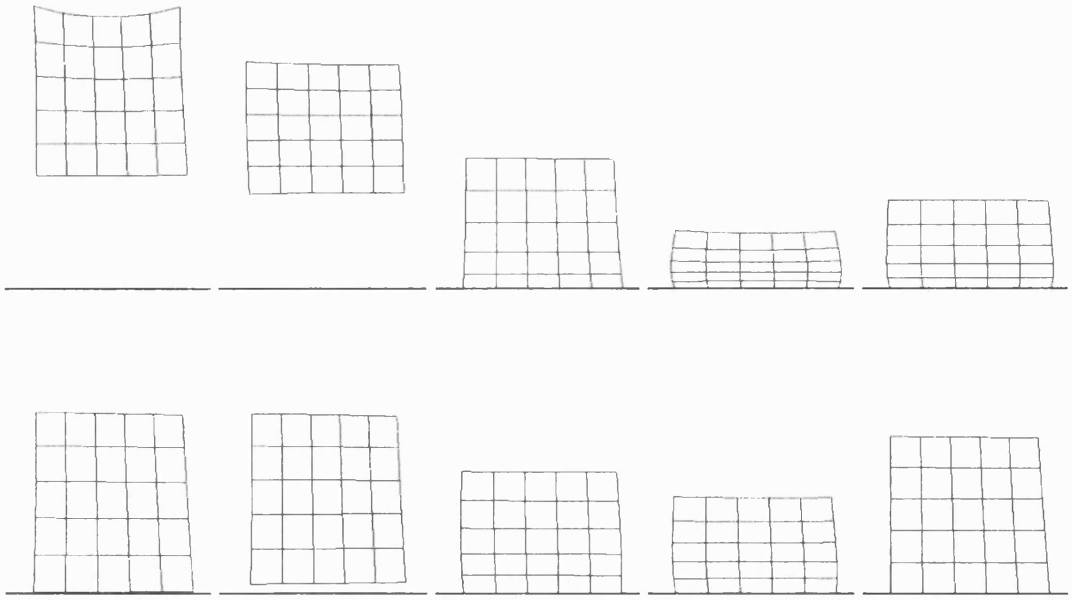


Figure 6.11: A solid suspended at two of its top corner nodes is released and it collides with the floor

We can see in this sequence how the solid gets squashed on impact with lateral bulging (due to the non-zero value of the Poisson ratio). It then rebounds and even lifts itself back over the floor. Eventually, the solid stabilises on the floor assuming a shape that is close to the original undeformed one.

Example 7: a solid suspended at only one of its top corner nodes is released and it collides with the floor

In this example, the setup is similar to the previous Example 6. We have a floppy solid but this time it is suspended at only one of its corner nodes. As the result, the simulation example starts with the solid hanging with the corner node opposite to the constrained one pointing directly downward (see the first frame on Figure 6.12 below).

It is an interesting example, because the solid falls with one of its corner nodes forward. Thus, when it collides with the floor, high level of strain will ensue in the area of the contact, which is just around the node in question.

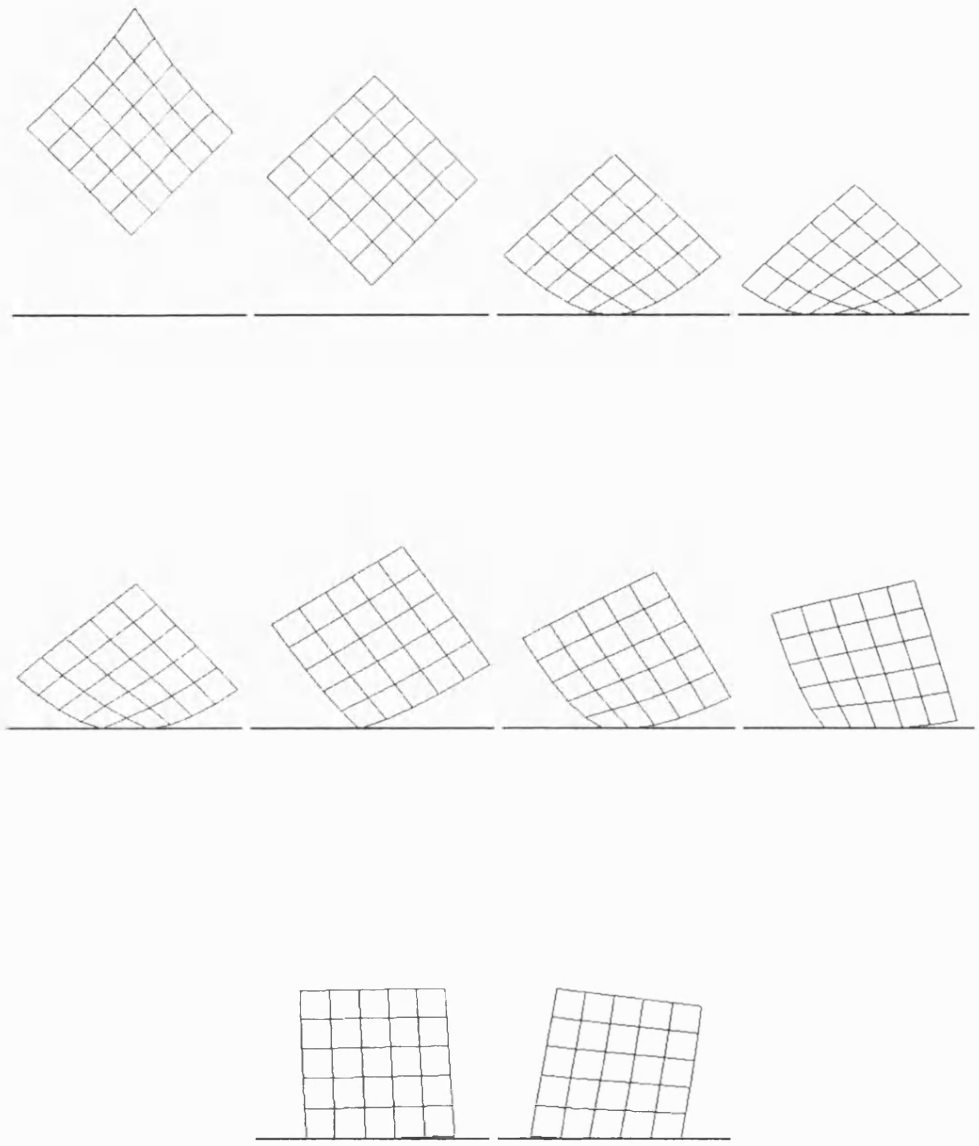


Figure 6.12: A solid suspended at only one of its top corner nodes is released and it collides with the floor

We can see that the results of the simulation look very realistic. The solid bounces off and falls on its side, where it finally settles down.

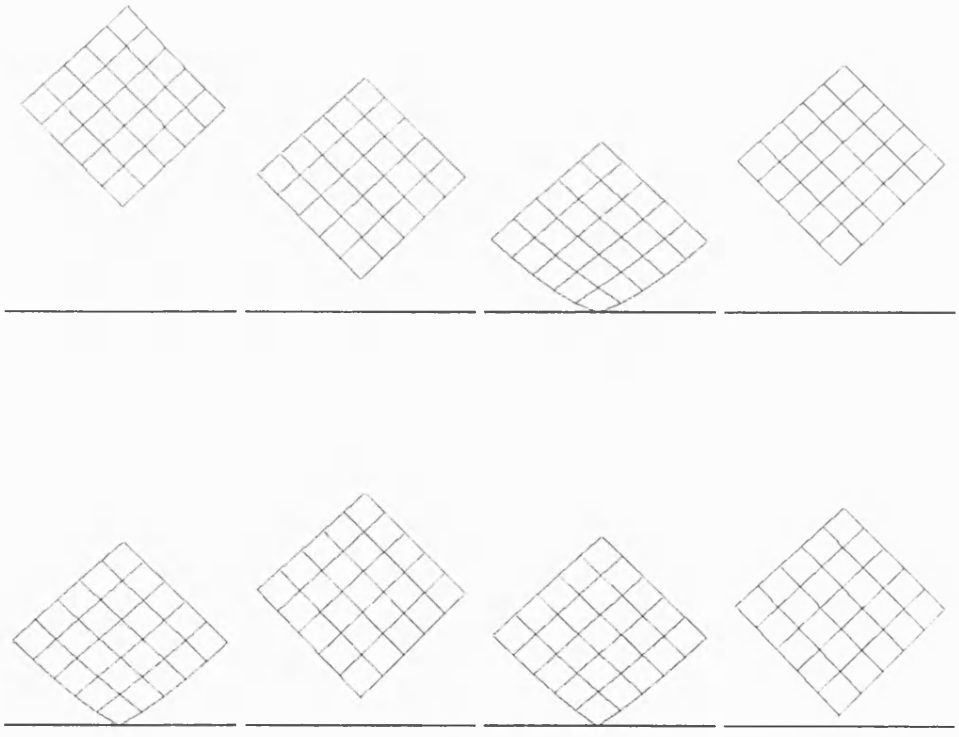
Example 8: a quasi-rigid solid suspended at only one of its top corner nodes is released and it collides with the floor

The following example is similar to the previous one but this time the solid is quasi-rigid (almost rigid). It is still elastic but is very firm, that is, with a very high value of the modulus of elasticity.

Normally this is a difficult problem to handle for any applied modelling technique. Very high values of the modulus of elasticity normally lead to so-called "stiff" equations. Special numerical techniques must normally be applied to solve stiff equations of motion. Otherwise, with the standard numerical techniques, the solver is prone to stability problems (the model may "blow up").

In this example, we aim to illustrate that our theory of FlexyMatter is able to handle stiff equations very well using the standard numerical solution. This is achieved with the help of the internal dissipation term, a force-like term that we used to model internal energy loss during solid's deformation. This dissipation term also serves to smoothen internal stress (strain) distribution inside the solid, levelling any sudden increases in stress (strain).

The following Figure 6.13 provides the resulting sequence of frames.



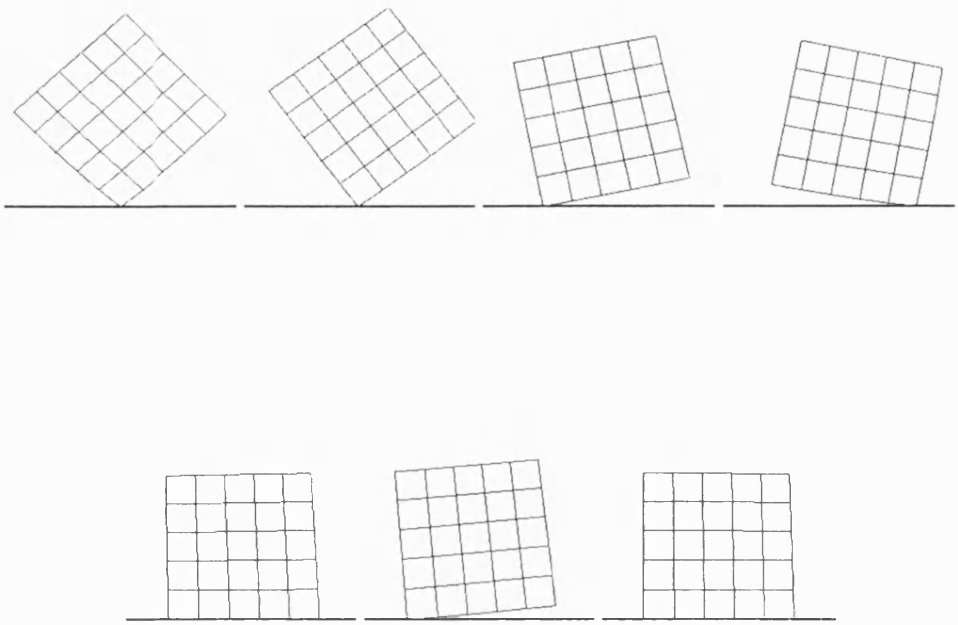


Figure 6.13: A quasi-rigid solid suspended at only one of its top corner nodes is released and it collides with the floor

We can visually confirm from this sequence that the theory has modelled our semi-rigid solid well. The solid falls on one of its corner nodes and bounces back twice with very little deformation. Eventually, after it settles down standing on the corner node, it falls on its side and comes to a stop after a couple of sideway swings.

We can see from this example that our theory of FlexyMatter can handle a wide range of elastic materials: from very floppy to semi-rigid. We can contrast this with other modelling techniques, in particular, with the method proposed by Terzopoulos et al [7, 12]. This modelling technique consists of two separate formulations, the primary formulation and the hybrid formulation, designed to handle solids with different elastic properties: the primary formulation for highly elastic solids and the hybrid formulation for rigid and semi-rigid solids. To model the last two of our examples, the animator will first have to decide which of the formulations to use. It will most probably be the primary formulation for Example 7 and the hybrid formulation for Example 8.

As we can see from the animation sequences, the theory of FlexyMatter is capable of handling both Example 7 and Example 8 within the same framework. We have demonstrated that the primary and hybrid formulations of Terzopoulos are both included

into our formulation. Thus, the theory of FlexyMatter has a smooth transition from flexible objects to rigid objects without the need for any special cases. This is a significant advance over all previous theories.

6.5 Summary

In this chapter, we presented a practical numerical solution of the modelling problem, described by our theory of FlexyMatter.

We began with the derivation of an approximate model for our precise mathematical model in section 6.1. Then we discussed the numerical solution of the differential equations governing the solid's evolution with respect to time in section 6.2. Finally, in section 6.3 we presented a practical implementation of the numerical solution, used to produce all the examples in this thesis.

In section 6.4 we provided a list of the results we have obtained using our test bed system, in which the theory of FlexyMatter is used to model the deformation behaviour of solids under the action of applied forces.

7. Conclusion

In this thesis we undertook research to develop a practical theory for modelling a wide range of elastic deformations: from almost rigid to very elastic. We created a theory that is internally complete and mathematically rigorous, but at the same time is flexible, versatile and easy to use. Its primary application is considered to be the area of computer modelling of elastic objects.

7.1 Classical Theory of Elasticity

The problem of modelling elastic media was rigorously addressed more than a hundred years ago in the 19^{th} century with the development of the classical theory of elasticity. Since then the classical theory has been successfully applied in various areas such as the propagation of sound and other elastic waves, structural design to assess the strength of constructions and so on. It has been shown to be a simple and relatively easy to use modelling theory and large amount of practical experience was accumulated.

The classical theory of elasticity provides a complete mathematical framework for modelling elastic deformations under the action of applied forces, including a mathematical description of solids, their material properties and applied forces. It also provides a system of differential equations and conditions describing the evolution of the solid for any given external forces and solid's material properties. However, as with most other mathematical theories, no explicit mechanism of the solution of these equations is available. Numerical techniques as well as a range of simplifications of the theory or special cases (such as linearized strain, simple explicit external forces, fixed solid's boundary and so on) had to be employed in order to find solutions of the equations for practical problems.

In the mid 1980s, with the rise of available computational power, the problem of practical physically based modelling of elastic solids started to attract large amount of interest from the research community. However, the linear classical theory of elasticity has serious practical limitations when applied directly to modelling highly elastic solids, and therefore it was never seriously used as a foundation of any practical modelling techniques. Surprisingly, no attempts were made by the research community in computer animation to extend the linear theory and remove its limitations or even simply investigate the limitations in detail.

7.2 Existing Modelling Techniques

Most of the practical modelling techniques developed from the mid 1980s (discussed in detail in chapter 3) use an energy-based approach. In this approach the solid to be modelled is prescribed a scalar energy – a non-negative scalar function of the points of the solid, which is the measure of its deformation. The energy grows when the solid is deformed and it has a minimum when the solid is in an equilibrium state. The modelling task is then a problem of finding the minimum of the solid's energy at any time during its evolution.

An energy-based approach is easy to understand and use, and it provides for fast numerical implementations. However, it has a number of serious limitations, including:

- It normally has a large number of parameters that need adjusting, normally on a trial and error basis.
- It provides no clear description for the boundary forces. This is a serious issue, since all interactions of a solid with other solids and external objects are performed through its boundary.
- It is only well suited for modelling highly elastic deformations. Modelling quasi-rigid solids is normally a difficult problem in practice. Some hybrid models were proposed to solve this problem (for example the hybrid formulation by Terzopoulos et al [**12**]), but they are complex and difficult to use in applications.
- More limitations and problems specific to actual energy-based modelling techniques are discussed in chapter 3.

We believe a completely different approach to modelling is required to solve the above given range of problems with the energy-based approach. In order to have a better picture of

what sort of approach is required, it is useful to list the attributes of a successful theory for modelling elastic solids:

- It should be rigorous. It should be based on the available classical theories (such as the classical theory of elasticity) as much as possible. We believe it is a mistake not to utilise the wealth of research and practical experience that was accumulated over the last 150 years.

- It should be complete. It should provide all the means necessary to account for all practical issues that arise in modelling elastic solids. This should include as a minimum surface forces and internal energy dissipation.

- It should be flexible, versatile and afford fast numerical implementations for use in near real-time animation. It should be able to deal with a wide range of elastic deformations within the same framework without the associated stability problems.

- It should provide a good foundation for future work in respect of creating a versatile framework for modelling complex elastic solids.

7.3 The theory of Hyper-Matter by Smith and Paddon

In 1992 Smith, Paddon [**89,90**] published their research findings on modelling elastic solids. They employed some pioneering ideas, not used before in the area of applied modelling, and developed a theory of Hyper-Matter as an extension of the classical theory of elasticity. It aimed specifically at resolving the limitations of the classical theory with regards to modelling highly flexible deformable objects. This was the first (and excepting our theory of FlexyMatter the only) theory that was developed as a direct extension of the classical theory of elasticity and not based on the energy approach.

Whilst in the linear classical theory of elasticity one static reference frame was used, Smith, Paddon suggested to use different reference frames at different times and for different points of the solid. Hyper-Matter was a three-dimensional object defined over a six-dimensional domain, which contained in itself the deforming solid as well as all of its reference frames at all times.

As the result, the theory of Hyper-Matter is a rigorous, versatile theory, based directly on the classical theory of elasticity and which is capable of modelling a wide range of deformations, far beyond the capabilities of the linear classical theory.

However, there are several reasons as to why the theory has not enjoyed a wide acceptance in the research community since 1992.

Firstly, and most importantly, it was written in a very terse and incomprehensible style. The authors seemingly were not interested in the popularisation of the theory. Very terse presentation was combined with few explanations, few graphical illustrations and complete lack of illustrative examples to present the fundamental ideas.

Secondly, it had a number of errors and inconsistencies. The most important of these, is the claim that in the theory of Hyper-Matter it is possible to use second partial derivative of the displacement field instead of second covariant derivative. While it is true that one can use the first partial derivative of the displacement field instead of the first covariant derivative, it is a mistake to do the same in the case of the second derivative. If this is done in curvilinear material coordinate systems, a force-like error term will be introduced, leading to seriously incorrect modelling results. This issue is discussed in detail in section 5..

And finally, the whole idea of Hyper-Matter, a three-dimensional object defined over a six-dimensional domain, in our opinion is questionable. As we show in this thesis, Hyper-Matter is an unnecessarily complex interpretation of a set of simple ideas. A much simpler theory can be derived, which preserves all the advances of the theory of Hyper-Matter over the other modelling techniques, but is much more accessible.

Smith, Paddon have introduced some pioneering ideas into the area of practical modelling of elastic solids, including the local alignment of the reference frame. We believe it is the actual interpretation of those ideas by Smith in his PhD thesis [**89**] that held back the popularisation of the theory of Hyper-Matter as a practical modelling technique.

Our intension was to use these fundamental ideas and develop a completely new theory with an emphasis on the simplicity of its presentation and the ease of use in applications, while still remaining firmly within the framework of the classical theory of elasticity. The theory of FlexyMatter we present in this thesis is the result of our research, which is an accessible and flexible, yet mathematically rigorous theory for modelling elastic deformations.

The next section provides a detailed overview of our research conclusions leading to the theory of FlexyMatter.

7.4 Our theory of FlexyMatter

Our theory of FlexyMatter is an extension of the linear classical theory of elasticity. Therefore the best way to present the theory, which we employed in this thesis, is first to present the classical theory of elasticity (see chapter 2), then undertake an investigation into why it is not suitable for modelling highly elastic solids and only then, when we had established the problematic areas of the linear classical theory, we introduce our proposed modifications. In this way, it is much easier to understand our theory and its fundamental ideas. We will use this approach now to give a brief outline of FlexyMatter.

7.4.1 Limitations of the linear classical theory of elasticity

In order to provide a better understanding of the way FlexyMatter extends the classical theory, it is useful to give an overview of the limitations of the linear classical theory, which make it unsuitable for modelling highly floppy solids.

We have shown (see section 4.2.1) that it is in fact the local rotational misalignment of the reference frame and the deforming body at a point that introduces errors in the evaluation of the stress tensor and therefore of the elastic force at the point. We have shown this by rotating the reference frame relative to the solid and examining the effect this rotation has on the first and second derivatives of the displacement field at a point (because these derivatives take part in the expressions for the strain tensor, the stress tensor and the elastic force).

We have also derived a specific condition for the misalignment expressed via the displacement field:

$$curl\,\mathbf{U} \neq 0 \tag{7.90}$$

We have shown that the curl of the displacement field $curl\,\mathbf{U}$ relates to the angle of the relative rotation of the local neighbourhood of a point on the reference frame relative to the one on the deforming solid.

This misalignment problem seriously limits the range of deformation the linear classical theory can handle. Indeed, for any deformation that results in the rotational misalignment of the reference frame and the deforming solid at some points of the solid, the elastic forces at those points will be evaluated with error. The error will depend directly on the degree of the

rotational misalignment. It is easy to see that even rigid body rotations without any actual deformation cannot be modelled in the classical theory.

7.4.2 Extension of the classical theory (theory of FlexyMatter)

Before we present our extensions, we note the following. The linear classical theory of elasticity suggests the following path for the evaluation of the elastic force at a point: we first calculate the displacement field at the point, then, using its derivatives, calculate the strain tensor, then the stress tensor and finally the elastic force. This procedure is purely local for the point in question. We only use the infinitesimal neighbourhood of the point to evaluate the relevant quantities. Therefore, if this local neighbourhood is rotationally misaligned, or mathematically speaking, $curl\,\mathbf{U} \neq 0$, the displacement field will contain a rotational component, which will lead to erroneous evaluation of all the quantities, starting with the strain tensor and finishing with the elastic force.

In the theory of FlexyMatter we have added another step into this classical path of evaluation of the elastic force: before calculating the displacement field in the neighbourhood of a point, the deforming solid and its reference frame should be aligned, so that $curl\mathbf{U} = 0$. Then the displacement field can be calculated, followed by the evaluation of all other related quantities (strain, stress and finally the elastic force). This time because we have done the local alignment and effectively removed the rotational component from the displacement field, the resulting elastic force will be free of the error due to the local rotational misalignment.

This is a fundamental idea, first introduced by Smith, Paddon in their theory of Hyper-Matter. It removes all the limitations of the classical theory of elasticity in one single stroke. It is therefore even more surprising to see that in fact the local rotational alignment simplifies the expression for the strain tensor.

Indeed, since:

$$curl\,\mathbf{U} = 0 \tag{7.91}$$

This can be used to show that the expression for the strain tensor is then simply:

$$\varepsilon_{ij} = U_i\big|_j \tag{7.92}$$

This simplification of the expression of the strain tensor leads to a very simple expression for the elastic force for the internal points:

$$E^i = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} U^i\Big|^j_j \tag{7.93}$$

If we, additionally to the rotational alignment, perform a transitional alignment as well, that is, we align in space the corresponding points on the deforming solid and the reference frame, we can show that this leads to further simplification of the formulation. In particular, in this case the first covariant derivative of the displacement field is equal to the ordinary first partial derivative. This simplifies further the classical expressions for the strain tensor, the stress tensor and the elastic force.

7.4.3 The theory of FlexyMatter as an advanced modelling technique

Our theory of FlexyMatter is a direct extension of the linear classical theory of elasticity and as such is mathematically rigorous. It does not compromise on the mathematical rigour to produce an easy to use modelling technique. It is based on the classical theory of elasticity and therefore benefits from its long established status and the wealth of the accumulated experience.

FlexyMatter is based on the same fundamental ideas as the theory of Hyper-Matter by Smith, Paddon, but presented in a much more accessible form. We believe that anybody with a good mathematical background should be able to understand our theory of FlexyMatter and create fast and flexible practical implementations in a relatively short period of time. Moreover, we suggested a number of practical improvements including:

- The procedure of incremental alignment of the reference frame and the deforming solid at a point. This procedure is used in practical discrete models and is based on the fact that at time t_{k+1} the reference frame for a point is only slightly different from the one at the previous time t_k. Therefore, the curl of the displacement field can be used to approximate the rotational angle, which can be used to rotate the reference frame to obtain the required alignment at time t_{k+1}.

- The development of the dynamic equations of motion for the boundary points, which include boundary forces in their formulation. This allows in practical discrete models to work with one single dynamic system of equations of motion for all points

of the solid, rather then having to deal with two different, coupled systems of dynamic and static equations for internal and boundary points.

We can therefore say that our theory of FlexyMatter effectively embraces and improves the theory of Hyper-Matter by Smith, Paddon. And the simplicity of our three-dimensional tensor system with 9 degrees of freedom contrasts well with the complexity of the six-dimensional tensor system from the theory of Hyper-Matter with its 81 degrees of freedom.

In section 6.4 we present practical examples of simulations using our theory of FlexyMatter. These examples show that indeed the theory of FlexyMatter is flexible and versatile. We have also shown that the theory can handle both very floppy and very rigid solids with ease. This is a significant advance over all previously proposed theories. The majority of these other modelling theories are energy based and are either unable to model quasi-rigid bodies at all or must provide an alternative formulation for this case (see for example the hybrid formulation by Terzopoulos et al [**12**] discussed in sections 3.1, 3.2 and 6.4). We can therefore again say that our theory of FlexyMatter embraces and improves other modelling theories.

Thus, our theory of FlexyMatter provides an excellent foundation for the future research and the development of a sophisticated modelling system capable of handling complex solids and their assemblies. In the next section we present a brief discussion of possible future work.

We have shown that our theory of FlexyMatter satisfies all the criteria for a successful modelling theory set out earlier. It is rigorous, complete, flexible and versatile. It embraces and extends the earlier work by Smith, Paddon (his theory of Hyper-Matter) as well as most of the other modelling theories developed over the last twenty years (in particular, the primary and hybrid formulations by Terzopoulos et al [**12**]).

7.5 Future work

There are several areas where the future research can expand or improve the theory of FlexyMatter. These include:

- Effective mechanism for handling complex solid boundaries. This will allow modelling solids with complex shapes.

- Handling constraints and constrained movement, in particular with respect to handling complex assemblies of solids. This problem is closely related with the next item.

- Object collision and self-penetration detection and response. Good practical algorithms and techniques need to be developed for object collision detection and response.

- The usage of other more efficient numerical solution techniques should be investigated. This is a practically important topic from the point of view of real-time animation.

Once most of these areas have been investigated, the theory of FlexyMatter can be further combined with the behavioural modelling systems in order to create a system capable of modelling genuinely complex objects, so that the animators can specify their behaviour and the system will automatically handle all the "routine" modelling aspects.

Appendix A: Overview of numerical methods

A.1 Explicit and implicit numerical schemes

A typical and simplest example of an explicit numerical scheme is the explicit Euler method. In this method, the system state at the next step is evaluated in a simple iteration:

$$\Lambda(l,m,n,t_{k+1}) = \Lambda(l,m,n,t_k) + \Delta t\,\Omega(\Lambda(o,p,q,t_k))$$

The expression on the right is a first order approximation of the first derivative in time. The current system state is used to calculate how to advance the system to the next step.

The time step in this simple example is constant, so that:

$$t_{k+1} = t_k + \Delta t, \forall k \geq 0$$

This explicit Euler method is of little practical value because all the shortcomings of explicit numerical methods listed at the end of this section are predominant in it. Other more advanced explicit methods, such as Runge-Kutta or Adams methods, should be used in any practical applications.

A simplest example of an implicit method is the backward or implicit Euler method:

$$\Lambda(l,m,n,t_{k+1}) = \Lambda(l,m,n,t_k) + \Delta t\,\Omega(\Lambda(o,p,q,t_{k+1}))$$

As we have mentioned in the brief description of the implicit methods, the system state at the next step $\Lambda(l,m,n,t_{k+1})$ is involved in either side of the equality. This system of

equations must be solved as a large system of linear algebraic equations for every time step in order to advance the system to the next step.

The implicit Euler method is more suited for practical applications, but still, as in the case of the explicit Euler method, more advanced implicit methods are recommended in practice. One advanced implicit numerical method that is suitable for our modelling problem is that the D'Yakanov's method used by Smith in his PhD thesis [**89**].

In general, explicit methods are faster, easier to implement and deal with, but also less precise and can have severe stability problems. The implicit methods, on the other hand, are generally more accurate, much more stable but also slower and more difficult to implement.

A.2 Numerical methods with time step control

Generally, methods that provide time step control, in addition to the description of the means to advance the system to the next instance in time, provide a means to estimate the error for the current time step Δt_k. Also, they provide a means to estimate the time step needed to achieve required accuracy.

More precisely, numerical methods with step control have means to evaluate the error $\Delta \varepsilon_k$ for the current time step Δt_k. Then, having estimated the error, there is also a means to evaluate the adjusted time step, which would achieve the required preset accuracy if it were used. Namely, if we denote the desired accuracy as A, the time step required to keep the error at the level of A is given by an expression, specific to the numerical scheme:

$$\Delta \bar{t}_k = f(\mathrm{A}, \Delta \varepsilon_k, \Delta t_k)$$

The estimation function $f(\mathrm{A}, \Delta \varepsilon_k, \Delta t_k)$ gives an estimate of the time step, which, if used, would produce an error equal or smaller than the given accuracy level A.

Having obtained the corrected time step, it can be used instead of the old one to advance the system to the next step. In this scenario, the global evaluation error can be kept at a constant level (preset beforehand), as the time step is adjusted (shortened or lengthened) according to the behaviour of the right part $\boldsymbol{\Omega}(\boldsymbol{\Lambda}(o, p, q, t_k))$.

A.3 Multi-step and single-step numerical methods

In single-step methods only the system state at the current step is used to evaluate the system state at the next step. Most of the practically available numerical methods (most notably Runga-Kutta and their variants) are single-step.

Multi-step numerical methods, on the other hand, use the states of the system at more than one previous step to evaluate the system state at the next step. Adams methods are the most widely used class of methods of this kind:

$$\Lambda(l,m,n,t_{k+1}) = \Lambda(l,m,n,t_k) +$$

$$+\frac{\Delta t}{12}(23\cdot\Omega(\Lambda(o,p,q,t_k)) - 16\cdot\Omega(\Lambda(o,p,q,t_{k-1})) + 5\cdot\Omega(\Lambda(o,p,q,t_{k-2})))$$

Multi-step methods normally have an adjustable number of previous states it uses for its computation. Because at the initial step t_0 there are no previous system states, it has to start off using only the current system state.

Multi-step methods require fewer evaluations per one time step than one-step methods (because the system states can be reused in subsequent evaluations), but are normally slower than the one-step methods.

A.4 Higher order numerical methods

Higher order numerical methods use more terms of the Taylor expansion and therefore do more intermediate evaluations per time step thus achieving higher accuracy. Examples are 4^{th} or 5^{th} order Runga-Kutta methods.

The disadvantage of these methods is obvious: more evaluations lead to a higher computational load and reduced speed of each step evaluation. A balance normally has to be found between the speed and accuracy (and stability).

A.5 Hybrid numerical methods

Hybrid methods are those that combine various specific features from different classes to produce methods best suited for particular applications.

In the case of our modelling problem, we need a numerical method that would satisfy the following requirements:

- It should be fast, in order to achieve interactive animation rates.
- The accuracy is not of the highest concern, but we do need to keep the error within some bounds.
- Stability is paramount. The method must cope with extreme deformations of the solid without blowing up and becoming unstable.

In our opinion, the method that suits all of the above-mentioned requirements best is the 5^{th} order Runga-Kutta method with step control. It is one-step and explicit, thus relatively fast. It is quite accurate and it provides a support for error evaluation and time step adjustment, all of which make it quite stable and able to cope with extreme solid's deformations.

Appendix B: The theory of Hyper-Matter by Smith and Paddon

In 1992 Smith and Paddon published their pioneering work on modelling deformable solids in the form of a technical report [**90**] and Smith's PhD thesis [**89**], in which they developed the theory of Hyper-Matter as a direct extension of the linear classical theory of elasticity. The theory of Hyper-Matter appeared only a few years after the results of the research findings by Terzopoulos et al [**7**, **12**] were published, which are still regarded as the first comprehensive attempt to develop an applied theory for modelling deformable solids.

In the theory of Hyper-Matter Smith and Paddon suggested expanding the three-dimensional model from the classical theory of elasticity into six-dimensions. They introduced a notion of a hyper-body given by the position vector $\mathbf{H}(x,y,t)$, which is a three-dimensional body defined over a six-dimensional domain, as an extension to the classical notion of the reference frame. This hyper-body, or the extended reference frame, contains both the deforming solid itself and infinitely many copies of the reference frames used to calculate elastic forces at individual points. To help define $\mathbf{H}(x,y,t)$ unambiguously they also introduced three so-called reference conditions, the second and the third of which relate to translational and rotational alignment of the reference frame and the deforming solid at each point of the solid.

This approach suggested by Smith and Paddon is not entirely new. The idea of local rotational alignments between the reference frame and the deforming body (or rather fluid) to increase applicability of linear theory is widely used in fluid mechanics (see for example corotational models in [97]). Various special cases were also considered in the theory of elasticity (see for example section 14 in [96]) in an attempt to deal with the local rotational

misalignments. But Smith and Paddon were first to successfully apply the ideas from the fluid mechanics (see corotational models in [97]) to modelling solids and obtain a theory that is capable of modelling such a rich variety of deformations within the linear framework.

Despite its novel approach, the theory of Hyper-Matter was poorly received. It was presented in a very terse and almost incomprehensible style. It had no examples and no illustrations, and in many cases the ideas were not fully developed to a form suitable for a direct implementation. In general, it did not look like a complete theory but rather like a collection of ideas presented in a very dry tensor format.

In this appendix we aim to explain the theory of Hyper-Matter with illustrations and clarifications. However, it is not an easy task. Due to extreme terseness of the original presentation, we had to add the interpretations and the description of the significance of some of the facts and ideas, which may of course differ in some detail from those of the authors. But the ultimate aim of this appendix is not to interpret the theory of Hyper-Matter (though we will in the conclusion express our views on it) but rather help an interested reader understand it faster and more easily. This will enable our own theory of FlexyMatter together with its historical counterpart, the theory of Hyper-Matter, to provide a complete description of the alternative approach to modelling elastic solids based on a direct expansion of the classical theory of elasticity. This should help an interested party to spend minimal time on understanding the basics of the approach, presented effectively in two different forms, and start working on improving and developing further our theory of FlexyMatter (and ultimately the theory of Hyper-Matter).

B.1 Overview of the theory of Hyper-Matter

We will now provide a detailed overview of the theory of Hyper-Matter. The structure of this section mirrors the structure of Smith's PhD thesis [**89**]. For each of the major chapters (all except for the introduction and the conclusion) we provide a description of all the key facts with examples and illustrations together with our own background explanations and discussions whenever we felt necessary. This section together with Smith's PhD thesis and a classical text on the theory of elasticity (for example Flugge [**91**]) should help an interested reader get a good understanding of the theory of Hyper-Matter developed by Smith and Paddon [**90**] and presented formally in Smith's PhD thesis [**89**].

B.1.1 The classical theory of elasticity

In this chapter the key facts from the theory of elasticity are presented. But first we introduce the notation used in this chapter.

We have a global Cartesian coordinate system with the base vectors $\boldsymbol{\delta}_k$. We also assume that our solid to be modelled is given by a vector function

$$\mathbf{R}(x^k): \mathrm{B} \to \mathbf{R}^3 \tag{B.1}$$

Where $\mathrm{B} = \{x \in \mathbf{R}^3 : |x^i| \le 1, i = 1,2,3\}$ is the three-dimensional unit cube.

As the solid deforms, it is represented by

$$\mathbf{r}(x^k, t): \mathrm{B} \to \mathbf{R}^3 \tag{B.2}$$

We have of course $\mathbf{r}(x^k, 0) = \mathbf{R}(x^k)$.

The curvilinear coordinate systems on the solid in its deformed and undeformed states are introduced as usual:

$$\mathbf{G}_i(x^k, t) = \frac{\partial \mathbf{r}(x^k, t)}{\partial x^i}$$
$$\mathbf{g}_i(x^k) = \frac{\partial \mathbf{R}(x^k)}{\partial x^i} \tag{B.3}$$

The dependency of all the quantities on x^k is often assumed but not formally identified. Also, instead of using the traditional tensor notation of dependency of one tensor or vector quantity on another, for example: $A_{ij} = A_{ij}(x^k)$, the following notation is used $A_{ij} = A_{ij}(x)$. Here x is always a three-dimensional vector and should be understood as the traditional x^k.

Smith introduces his own notation for the covariant derivative of a tensor. He uses the semicolon instead of the usual vertical bar: $U_{j;i} \equiv U_j|_i$ as in, for example, Flugge [**91**].

As the present author points out himself, this chapter on the classical theory of elasticity is taken largely from Flugge [**91**]. We will therefore only present a brief summary of the key facts needed in this chapter.

Cauchy's formula

According to Cauchy's formula the elastic force acting through an arbitrary element of area at any point in the solid can be represented with the aid of a symmetric tensor of a second order σ^{ij} defined at that point, called the stress tensor. Namely for any area element represented by a vector $\mathbf{dS} = dS_i \mathbf{g}^i$, the elastic force transmitted through this area element can expressed via the stress tensor $dF^i = \sigma^{ij} dS_j$.

Cauchy's theorem

Cauchy's theorem states that if we split the solid into two volumes V_1 and V_2 with a common boundary ΔS, the elastic force $\mathbf{T}_{(1)}$ due to V_1 acting on V_2 through ΔS is equal in length and opposite in direction to the force $\mathbf{T}_{(2)}$ due to V_2 acting on V_1 through the boundary ΔS. Namely:

$$\mathbf{T}_{(1)} = -\mathbf{T}_{(2)} \tag{B.4}$$

Using Cauchy's theorem Smith shows that the total sum of all internal elastic forces is always zero:

$$\int_V \mathcal{E}^i \mathbf{g}_i dm = 0 \tag{B.5}$$

Here $\mathcal{E}^i$ is the internal elastic force acting on the mass element dm.

The divergence theorem (Gauss)

Gauss's divergence theorem is well known. It links the circulation of a vector field along the boundary of an element of volume with the divergence of the field inside the element:

$$\oint_S \mathbf{v} \cdot \mathbf{dS} = \int_V div(\mathbf{v}) dV \tag{B.6}$$

Applying the divergence theorem to the expression of the elastic force acting on the infinitesimal volume of a point $d\mathcal{E}^i = \oint_S \sigma^{ij} dS_j$, we can obtain the following expressions:

$$d\mathcal{E}^i = \oint_S \sigma^{ij} dS_j = \sigma^{ij}_{;j} dV$$

$$\mathcal{E}^i = \sigma^{ij}_{;j} \tag{B.7}$$

Where $\mathcal{E}^i$ is the elastic force measured in force for unit volume.

Strain tensor

In this section the strain tensor ε_{ij} is introduced as a measure of the solid's deformation:

$$\varepsilon_{ij} = \frac{1}{2}\left(G_{ij} - g_{ij}\right) \tag{B.8}$$

If we now introduce the displacement field, as usual:

$$\mathbf{U} = \mathbf{r} - \mathbf{R} \tag{B.9}$$

We can express the strain tensor in terms of the displacement field (again, this is a classical expression):

$$\varepsilon_{ij} = \frac{1}{2}\left(U_{i;j} + U_{j;i} + U^{k}_{;i} U_{k;j}\right) \tag{B.10}$$

This non-linear expression is the full expression for the strain tensor. As in the classical case, Smith suggests that it should be dropped (on the assumption of small gradients $U_{j;i}$) and the linear expression should be used instead:

$$\varepsilon_{ij} = \frac{1}{2}\left(U_{i;j} + U_{j;i}\right) \tag{B.11}$$

Constitutive equations

The classical expression in the form of the Hooke's law relating the stress and the strain tensors is used:

$$\sigma^{ij} = E^{ijkl}\varepsilon_{kl} \tag{B.12}$$

Where E^{ijkl} are the elastic moduli.

Assuming that the solid to be modelled is homogeneous and isotropic, it can be shown that of total 81 components of E^{ijkl}, there are only two that are independent. Smith uses the Lame's constants λ and μ, so that the stress-strain relationship is simply:

$$\sigma^{ij} = \left(\lambda g^{ij} g^{kl} + \mu\left(g^{ik} g^{jl} + g^{il} g^{jk}\right)\right)\varepsilon_{kl} \tag{B.13}$$

The Lame's constants are not normally used in engineering, as they are more convenient for theoretical manipulations. In actual applications it is more convenient to use the modulus of elasticity E and the Poisson ratio ν due to their clear physical meaning (see section 2.7.4 as

well as (2.239) for the formulae expressing Lame's constants in terms of the modulus of elasticity and the Poisson ratio).

The dynamics of deformable bodies

By this point, all the major facts from the classical theory of elasticity have been introduced and the equations of motion can be formulated. Smith, however, does not explicitly do so in his thesis. Instead, in this section he gives a short discussion about the choice of the position of the reference frame. But, before we present this discussion, we feel it is useful to present the equations of motion in two different forms, and only then follow with the discussion of the choice of the reference frame position.

First of all, the classical equation of motion for the internal points of the solid in terms of the displacement field is (see Flugge [**91**]):

$$\rho \ddot{U}^i = \sigma^{ij}_{;j} + F^i \tag{B.14}$$

Here F^i is the resultant of the external volume forces acting on the current point and all the quantities are expressed in the coordinate system $\mathbf{g}_i$ (on the reference frame).

There is also a static condition on the boundary, which balances the internal elastic forces on the boundary (stress) and the external boundary forces.

$$\sigma^{ij} n_j = S^i \tag{B.15}$$

Here n_j is the external unit normal at the current boundary point and S^i is the external boundary force.

The stress depends on the displacement field through the constitutive equations (more precisely on the covariant derivatives of the displacement field). Using (B.9) we can rewrite both the equation of motion (B.14) and the boundary condition (B.15) in a different form:

$$\begin{aligned} \rho \ddot{\mathbf{r}} &= \sigma^{ij}_{;j} \mathbf{g}_i + \mathbf{F} \\ \sigma^{ij} n_j \mathbf{g}_i &= \mathbf{S} \end{aligned} \tag{B.16}$$

Here $\mathbf{r}$ is the position vector of the current point of the solid.

In this formulation the dependency on the displacement field is still present but is confined only to the expression for the internal elastic force term $\sigma^{ij}_{;j}\mathbf{g}_i$ and the boundary elastic force $\sigma^{ij}n_j\mathbf{g}_i$. There are several notes we can make in this regard:

- If, for whatever reason, the position of the reference frame is changed, it can (and will) only affect the calculation of the elastic forces.
- The reference frame can be in different positions for each point of the solid at any time $t > 0$.
- Even if displaced, the reference frame must possess the same metric properties as the solid at the start of the experiment, or more precisely to have the same metric properties as the solid in the undeformed shape.

Classically, the reference frame is chosen to be the solid itself at the start of the experiment and this frame remains static throughout the experiment. This is, however, not the optimal choice; it can (and should) be chosen differently if we are to model large global deformations. Indeed, as we have shown in this thesis (see section 4.2) the classical expression of the elastic force via the covariant derivatives of the displacement field is very sensitive to rotational misalignment of the deforming solid and the reference frame. In the classical theory of elasticity small global deformations are normally assumed, which leads to small rotational misalignments, producing an acceptably small error in the calculation of the elastic force. When modelling globally large deformations, these errors due to the rotational misalignment become too large, leading to serious modelling errors.

This freedom of choice for the position of the reference frame when calculating the elastic force is the key argument used by Smith in this section to justify expanding the classical formulation into six dimensions.

B.1.2 A six-dimensional model for globally small deformations

In this section we present a background discussion, which should help readers understand the six-dimensional theory more easily. We then follow with the formal presentation of the six-dimensional theory for modelling globally large deformations, as presented in Smith's thesis. In the next section we will present the theory of Hyper-Matter itself, which, as we will show, is the special case of the six-dimensional theory presented in this section.

As we have already noted in the previous section, we have a freedom of choice of the reference frame when calculating the internal elastic force as a point of the solid. We can, for example, have a static reference frame that is the solid itself at the start of the experiment; this is the classical case, the same reference frame is used to calculate elastic forces at all times and for all solid points. We can, on the other hand, at each time step align the deforming solid and the reference frame globally (that is, align the centres of mass and average rotations of the deforming solid and the reference frame) and use this aligned reference frame to calculate the elastic forces at all points of the solid at that time step. We can continue along this line of thought and suppose that it is possible to find a procedure of local alignments, which uses a different reference frame for different points of the solid at different times. We have not determined how this can be done, but we do know at this stage that it is possible.

Earlier, we denoted the reference frame as $\mathbf{R}(x)$. This corresponds to the classical case when the reference frame is static. If we suppose that global alignment is used, it will become $\mathbf{R}(x,t)$ – different at each instant of time t. If we further assume the reference frame is different for each point of the solid, we need to use the following notation: $\mathbf{R}(x,y,t)$. Here, to be consistent with Smith's notation, we have renamed x to y and added a new variable x to denote the point on the solid for which this reference frame was calculated. We will, for the moment, call $\mathbf{R}(x,y,t)$ the generalised reference frame (or "shape" as favoured by Smith).

We can note a number of properties of $\mathbf{R}(x,y,t)$ due to the way we have constructed it:

- $\mathbf{R}(x,y,t)$ is a vector function defined over the Cartesian square product of the domain B: $\mathbf{R}(x,y,t):\mathrm{B}^2=\mathrm{B}\times\mathrm{B}\to\mathbf{R}^3$. It therefore still describes a three-dimensional object, which is defined over a six-dimensional domain. In other words, it maps the six-dimensional unit cube B^2 into the conventional three-dimensional space, producing, in general, a non one-to-one relationship between B^2 and the image of $\mathbf{R}(x,y,t)$ (a three-dimensional object).

- For any fixed x_0, the set $\{\mathbf{R}(x_0,y,t): y\in\mathrm{B}, t\geq 0\}$ is a rotated and displaced copy of the set $\{\mathbf{R}(y): y\in\mathrm{B}\}$. Or, mathematically speaking, the metric tensors on the two sets should be equal:

$$\frac{\partial \mathbf{R}(x_0,y,t)}{\partial y^i}\cdot\frac{\partial \mathbf{R}(x_0,y,t)}{\partial y^j}=\frac{\partial \mathbf{R}(y)}{\partial y^i}\cdot\frac{\partial \mathbf{R}(y)}{\partial y^j} \qquad \text{(B.17)}$$

- The actual alignment procedure is yet to be defined. We will show later that the alignment procedure is well defined and makes good physical sense.

We can now give several illustrations to aid understanding of what $\mathbf{R}(x,y,t)$ is in specific examples. First of all we note that the best way to provide illustrations for the theory of Hyper-Matter is to use pseudo one-dimensional objects (the objects are still three-dimensional but very thin in two of their dimensions).

Let us suppose we have a solid, which deforms during its evolution as shown in Figure B.1 below.

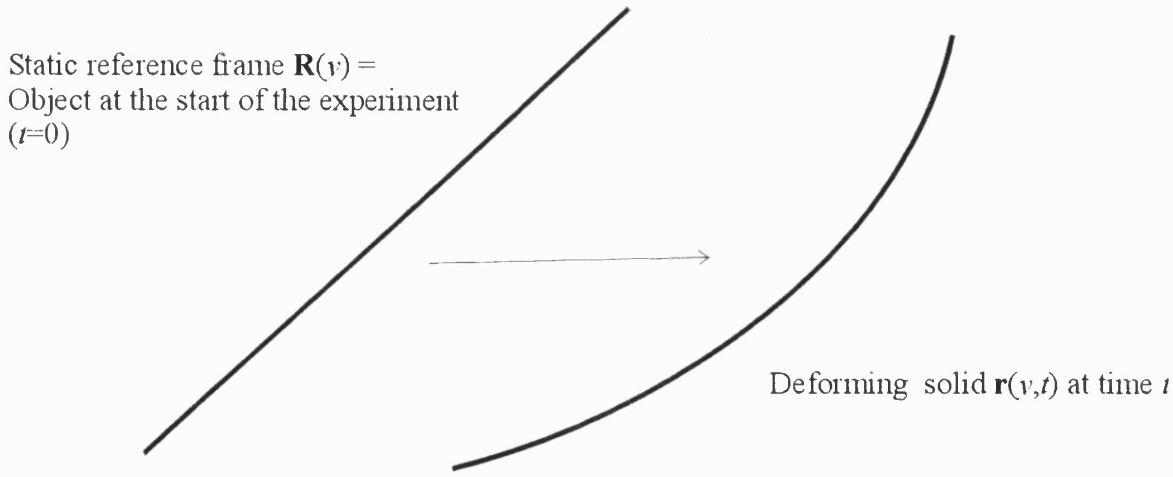


Figure B.1: Solid deformation with time

We now give an illustration of what $\mathbf{R}(x,y,t)$ may look like in this example (See Figure B.2).

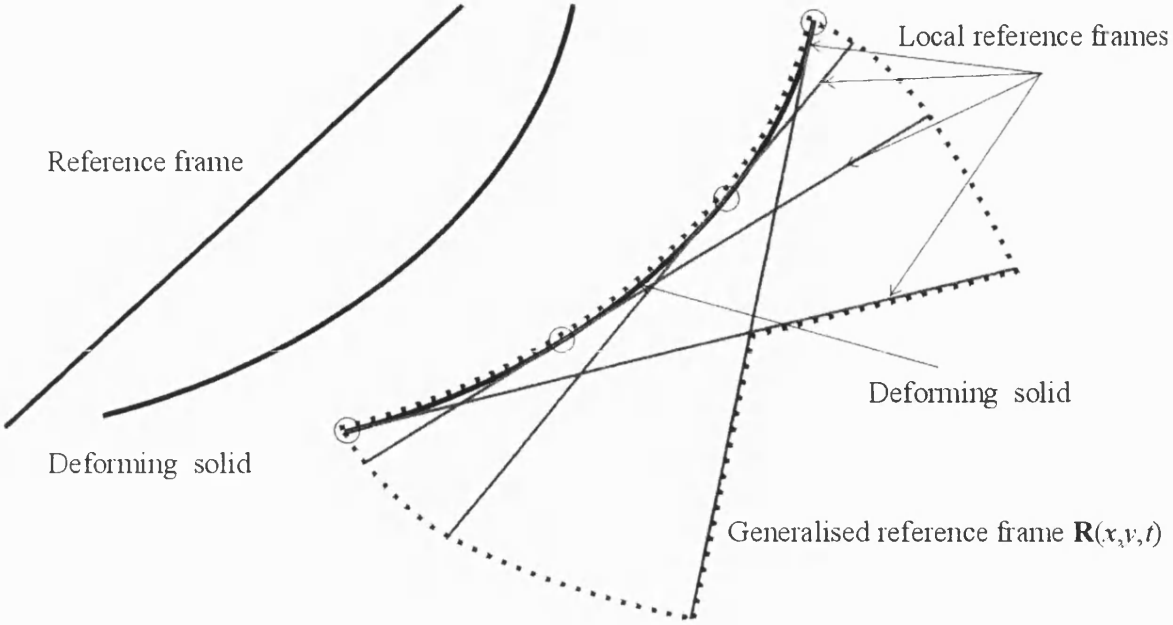


Figure B.2: Reference frame $\mathbf{R}(x)$, deforming solid $\mathbf{r}(x,t)$ and the generalised reference frame $\mathbf{R}(x,y,t)$

In this figure we have shown the reference frame $\mathbf{R}(x)$ and the deforming solid $\mathbf{r}(x,t)$ separately for convenience. The set $\{\mathbf{R}(x,y,t): x,y \in \mathrm{B}\}$ is outlined with a dotted line.

One can see from this illustration how the generalised reference frame is formed. For every point x_0 on the deforming solid, we align the set $\{\mathbf{R}(x_0,y,t): y \in \mathrm{B}\}$ so that:

- $\mathbf{R}(x_0,x_0,t)=\mathbf{r}(x_0,t)$. This is the translational alignment; we moved a copy $\{\mathbf{R}(x_0,y,t): y \in \mathrm{B}\}$ of the reference frame $\mathbf{R}(x)$ so that $\mathbf{R}(x_0,x_0,t)=\mathbf{r}(x_0,t)$.
- The translated copy $\{\mathbf{R}(x_0,y,t): y \in \mathrm{B}\}$ of the reference frame $\mathbf{R}(x)$ forms a tangent to the deforming solid at the point x_0. The exact mathematical expression for this rotational alignment will be given later, in the next section.

Let us now look at two sets: $\{\mathbf{R}(x,z_0,t): x \in \mathrm{B}\}$ and $\{\mathbf{R}(z_0,y,t): y \in \mathrm{B}\}$, where $z_0 \in \mathrm{B}$ (see Figure B.3 below). This will give further insight into the structure of the generalised reference frame $\mathbf{R}(x,y,t)$.

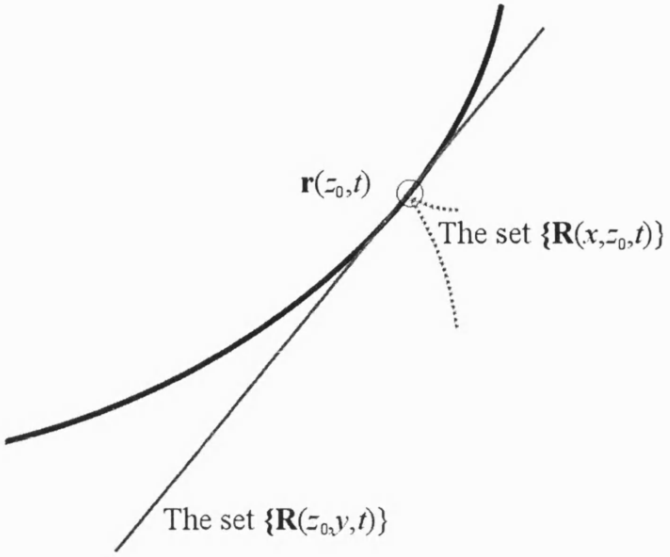


Figure B.3: Two sets $\{\mathbf{R}(x,z_0,t): x \in \mathrm{B}\}$ and $\{\mathbf{R}(z_0,y,t): y \in \mathrm{B}\}$ at a point $z_0 \in \mathrm{B}$

Indeed, as we have noted previously from the way we constructed $\mathbf{R}(x,y,t)$, the set $\{\mathbf{R}(z_0,y,t): y \in \mathrm{B}\}$ is an aligned (translated and rotated) copy $\mathbf{R}'(x)$ of the reference frame $\mathbf{R}(x)$ and the set $\{\mathbf{R}(x,z_0,t): x \in \mathrm{B}\}$ is just the path of $\mathbf{R}(x)$ as the copies are re-aligned at different points $x \in \mathrm{B}$.

From the examples and the way we constructed the generalised reference frame $\mathbf{R}(x,y,t)$, we can see that at all times it contains in itself the deforming solid and all the local reference frames (which are translated and rotated copies of $\mathbf{R}(x)$). So for any point $z_0 \in \mathrm{B}$, $\mathbf{R}(z_0,z_0,t)$ is the position vector of the point on the deforming solid (as well as on the aligned local reference frame represented by the set $\{\mathbf{R}(z_0,y,t): y \in \mathrm{B}\}$, because they are translationally aligned) and $\{\mathbf{R}(z_0,y,t): y \in \mathrm{B}\}$ is the locally aligned reference frame, which can be used to calculate the displacement field and all of its derivatives at point z_0 to be used in the calculation of the elastic forces at the point. Formally, $\mathbf{R}(x,y,t)$ is all we need to model the behaviour of the solid at all times. If we know $\mathbf{R}(x,y,t)$ at all times, the deforming solid is then simply described by the set $\{\mathbf{R}(x,x,t): x \in \mathrm{B}\}$.

The theory of Hyper-Matter has this notion of the six-dimensional (generalised) reference frame $\mathbf{R}(x,y,t)$ at its heart. It provides the necessary formalism in order to find $\mathbf{R}(x,y,t)$ for any given solid $\mathbf{R}(x)$ at the start of the experiment and any boundary and volume forces given at all times.

As we have seen, using a six-dimensional domain has its formal benefits. We have one single entity $\mathbf{R}(x,y,t)$ that is constructed in a specific way for any solid and has everything we need to model the deformation of that solid through time under the action of applied forces. But these formal simplifications come at a price. We will need to add the conditions for $\mathbf{R}(x,y,t)$ (so called reference conditions), which must be satisfied by the generalised reference frame $\mathbf{R}(x,y,t)$ at all times. Additionally, we will need to explain how the classical expressions of strain, stress and the elastic force are formulated in terms of this new object, defined over a six-dimensional domain. All the standard mathematical operations such as taking a derivative will need to be re-defined (or at least explained) and new operations such as taking a collateral derivative will need to be introduced.

The above introduction should be sufficient to understand the key facts from the relevant sections of Smith's thesis, presented below.

Construction of the reference shape

We first introduce important notation and definitions.

We will from now on use the following notation for the undeformed solid:

$$\mathbf{N}(x)=\mathbf{r}(x,0)=\mathbf{R}(x) \tag{B.18}$$

Natural base vectors are the base vectors on $\mathbf{N}(x)$:

$$\tilde{\mathbf{g}}_i(x)=\frac{\partial \mathbf{N}(x)}{\partial x^i} \tag{B.19}$$

The natural metric is the metric on $\mathbf{N}(x)$:

$$\tilde{g}_{ij}(x)=\tilde{\mathbf{g}}_i(x)\cdot\tilde{\mathbf{g}}_j(x) \tag{B.20}$$

For any vector quantity $\mathbf{A}(x,y,t)$, its derivative is defined as follows:

$$\mathbf{A}_{,i}(x,y,t)=\frac{\partial \mathbf{A}(x,y,t)}{\partial y^i} \tag{B.21}$$

That is, we differentiate with respect to the second variable y while holding the first one x fixed.

The delta neighbourhood $\delta V(x,y)$ of a six-dimensional point (x,y) is the following set of points in the domain B^2:

$$\delta V(x,y)=\left\{(x,z)\in \mathrm{B}^2 : \forall f \in \{1,2,3\}\, y^f-\delta \le z^f \le y^f+\delta\right\} \tag{B.22}$$

In addition to $\delta V(x,y)$, we will also introduce a set $\delta V(y)$, which is a three-dimensional analogue of $\delta V(x,y)$:

$$\delta V(y)=\left\{z\in \mathrm{B} : \forall f \in \{1,2,3\}\, y^f-\delta \le z^f \le y^f+\delta\right\} \tag{B.23}$$

We now define the mass element $dm(y)$:

$$dm(y)=\rho(y)\sqrt{\tilde{g}(y)}dy^1dy^2dy^3 \tag{B.24}$$

Where $\tilde{g}(y)=\det(\tilde{g}_{ij}(y))$.

We can now present the three reference conditions. These conditions are important for constructing a unique generalised reference frame $\mathbf{R}(x,y,t)$ for any particular solid with the natural shape $\mathbf{N}(x)$ and which is deforming as described by the mapping $\mathbf{r}(x,t), t\ge 0$ (of course $\mathbf{r}(x,0)=\mathbf{N}(x)$).

A mapping $\mathbf{R}(x,y,t):\mathrm{B}^2\to\mathbf{R}^3$ is a reference shape (or generalised reference frame as we used previously) for a solid with the natural shape $\mathbf{N}(x)$ and whose deformation is described by a mapping $\mathbf{r}(x,t), t\ge 0$ ($\mathbf{r}(x,0)=\mathbf{N}(x)$) if and only if $\mathbf{R}(x,y,t)$ satisfies the following three conditions:

- The first reference condition:

$$\mathbf{R}_i(x,y,t)\cdot\mathbf{R}_j(x,y,t)=\mathbf{N}_i(y)\cdot\mathbf{N}_j(y) \tag{B.25}$$

- The second reference condition:

$$\int_{\delta V(x)}\mathbf{R}(x,z,t)dm(z)=\int_{\delta V(x)}\mathbf{r}(z,t)dm \tag{B.26}$$

- The third reference condition:

$$\int_{\delta V(x)} \nabla \times \mathbf{R}(x,z,t)dm(z) = \int_{\delta V(x)} \nabla \times \mathbf{r}(z,t)dm(z) \qquad \text{(B.27)}$$

Here ∇ is the curl operator ($\nabla \equiv curl$).

We must note here that in the second and third conditions above Smith used $\mathbf{r}(x,z,t)$ instead of $\mathbf{r}(z,t)$, where $\mathbf{r}(x,y,t) \equiv \mathbf{r}(y,t)$. We believe it is quite unnecessary to extend the distinctive three-dimensional quantity $\mathbf{r}(x,t)$ into six dimensions. It seems appropriate to keep it three-dimensional, similarly to the natural shape $\mathbf{N}(x)$, and define the six-dimensional reference shape $\mathbf{R}(x,y,t)$ around these three-dimensional quantities ($\mathbf{N}(x)$ and $\mathbf{r}(x,t)$).

The first reference condition states simply that with a fixed first variable x, the set $\{\mathbf{R}(x,y,t): x - fixed\}$ is a rigid body displacement (a rotated and translated copy) of the natural shape $\mathbf{N}(x)$. The second and the third conditions are the alignment conditions; the second condition describes translational alignment (alignment of the centres of mass of the corresponding sets $\{\mathbf{R}(x,z,t): z \in \delta V(x)\}$ and $\{\mathbf{r}(x,z,t): z \in \delta V(x)\}$) and the third condition describes rotational alignment (alignment of the relative rotations of the sets).

The second and third conditions are formulated for the delta-neighbourhood of point x, not for the point itself (as in our examples above). This is simply a generalisation of the point case: instead of considering alignment conditions as equalities at specific points, we consider them as equalities of averaged quantities (presented as volume integrals). Despite the fact that δ can take any non-zero value, only two special cases present practical interest: when $\delta \to 0$ and when $\delta = 1$. We will discuss this issue in detail when we discuss the domain of linearity.

Smith does not expand on how the second and third conditions are obtained; he simply introduces them without much discussion. However, in this thesis we have provided all the necessary background on how these conditions can be obtained (for the limiting case when $\delta \to 0$, which is the case for Hyper-Matter) and why: see sections 4.2, 4.3 and additionally section 6.3.3, where we discuss practical alignment of the deforming solid and the reference frame.

We will now present the second and third reference conditions in a different form, using the generalised displacement field.

We first introduce the (generalised) displacement field:

$$\mathbf{U}(x,y,t)=\mathbf{r}(y,t)-\mathbf{R}(x,y,t) \tag{B.28}$$

Again, Smith used $\mathbf{r}(x,y,t)$ - a six-dimensional version of $\mathbf{r}(y,t)$, whilst we have avoided it for clarity (as before).

The second and third reference conditions can then be presented as follows:

$$\begin{aligned}&\int_{\delta V(x)} U^i(x,z,t)\mathbf{g}_i(x,z,t)dm(z)=\mathbf{0}\\&\int_{\delta V(x)} \in^{ijk} U_{j;k}(x,z,t)\mathbf{g}_i(x,z,t)dm(z)=\mathbf{0}\end{aligned} \tag{B.29}$$

Where $\mathbf{g}_i(x,z,t)=\mathbf{R}_{,i}(x,y,t)$ is the (generalised) basis on the reference shape.

Existence and uniqueness of the reference shape

In this section, Smith uses induction over time to show that the reference frame exists and that it is unique.

However, before we present the proof, we introduce several important facts, which are used in the proof and which will help in understanding it. First of all, we recall that the reference shape $\mathbf{R}(x,y,t)$ contains aligned copies of the natural shape $\mathbf{N}(x)$ at all points x and all times t. For any fixed point x and time t, alignment refers to the alignment of the two sets $\{\mathbf{R}(x,y,t): y\in \mathrm{B}\}$ and $\{\mathbf{r}(z,t): z\in \mathrm{B}\}$ and consists of two independent components: translational alignment (reference condition two) and rotational alignment (reference condition three).

Let us suppose the two sets $\{\mathbf{R}(x,y,t): y\in \mathrm{B}\}$ and $\{\mathbf{r}(z,t): z\in \mathrm{B}\}$ are misaligned at point x and both reference conditions two and three do not hold. In terms of the displacement field, this implies:

$$\begin{aligned}&\int_{\delta V(x)} U^i(x,z,t)\mathbf{g}_i(x,z,t)dm(z)=\overline{\mathbf{C}}(x)\neq\mathbf{0}\\&\int_{\delta V(x)} \in^{ijk} U_{j;k}(x,z,t)\mathbf{g}_i(x,z,t)dm(z)=\overline{\mathbf{\Theta}}(x)\neq\mathbf{0}\end{aligned} \tag{B.30}$$

These two vectors, $\overline{\mathbf{C}}(x)$ and $\overline{\mathbf{\Theta}}(x)$, are the measure of the misalignment; if they are both zero, the alignment conditions hold. If either one of them is not zero (or they are both not

zero) the second and/or third reference conditions are violated and $\{\mathbf{R}(x,y,t): y \in \mathrm{B}\}$ and $\{\mathbf{r}(z,t): z \in \mathrm{B}\}$ are misaligned either translationally or rotationally or both.

Let us look at these two quantities $\overline{\mathbf{C}}(x)$ and $\overline{\mathbf{\Theta}}(x)$ in more detail. $\frac{1}{M_{\delta V(x)}}\overline{\mathbf{C}}(x)$ is physically the average distance between the sets $\{\mathbf{R}(x,y,t): (x,y) \in \delta V(x,x)\}$ and $\{\mathbf{r}(z,t): z \in \delta V(x)\}$ (here $M_{\delta V(x)}$ is the mass associated with $\delta V(x)$). We can use it to correct the reference shape:

$$\tilde{\mathbf{R}}(x,y,t) = \mathbf{R}(x,y,t) + \frac{1}{M_{\delta V(x)}}\overline{\mathbf{C}}(x) \tag{B.31}$$

We can now see that the second reference condition will hold for the corrected reference shape $\tilde{\mathbf{R}}(x,y,t)$. Indeed:

$$\tilde{\mathbf{U}}(x,y,t) = \mathbf{r}(y,t) - \tilde{\mathbf{R}}(x,y,t) = \mathbf{r}(y,t) - \mathbf{R}(x,y,t) - \frac{1}{M_{\delta V(x)}}\overline{\mathbf{C}}(x) \tag{B.32}$$

Therefore

$$\int_{\delta V(x)} \nabla \times \tilde{\mathbf{U}}(x,z,t)\, dm(z) = \int_{\delta V(x)} \nabla \times \left(\mathbf{r}(z,t) - \mathbf{R}(x,z,t) - \frac{1}{M_{\delta V(x)}}\overline{\mathbf{C}}(x) \right) dm(z) = \tag{B.33}$$

$$= \int_{\delta V(x)} \nabla \times (\mathbf{r}(z,t) - \mathbf{R}(x,z,t))\, dm(z) - \overline{\mathbf{C}}(x) = \mathbf{0}$$

With the other quantity $\overline{\mathbf{\Theta}}(x)$, the situation is similar: $\frac{1}{M_{\delta V(x)}}\overline{\mathbf{\Theta}}(x)$ (or $\frac{1}{2M_{\delta V(x)}}\overline{\mathbf{\Theta}}(x)$ to be precise) is the average angle, through which the set $\{\mathbf{R}(x,y,t): (x,y) \in \delta V(x,x)\}$ can be rotated to "correct" rotational misalignment. For more information on why $\frac{1}{2M_{\delta V(x)}}\overline{\mathbf{\Theta}}(x)$ can be considered an angle of rotation of the set $\{\mathbf{R}(x,y,t): (x,y) \in \delta V(x,x)\}$ relative to the set $\{\mathbf{r}(z,t): z \in \delta V(x)\}$, please see Flugge [91] (p.86) or in this thesis see section 6.3.3.

Therefore, we can rotate the set $\{\mathbf{R}(x,y,t): (x,y) \in \delta V(x,x)\}$ through the angle $\frac{1}{2M_{\delta V(x)}}\overline{\mathbf{\Theta}}(x)$, and we will obtain the "corrected" $\tilde{\mathbf{R}}(x,y,t)$. That is, the third reference

condition will hold for this aligned $\tilde{\mathbf{R}}(x,y,t)$ (we do not present a mathematical proof due to its cumbersomeness).

Applying both arguments for $\overline{\mathbf{C}}(x)$ and $\overline{\mathbf{\Theta}}(x)$ in turn, we can obtain the corrected reference shape $\tilde{\mathbf{R}}(x,y,t)$, for which both the second and the third reference conditions will hold.

We can now present Smith's arguments for the existence and uniqueness of the reference shape $\mathbf{R}(x,y,t)$ for any deforming solid $\mathbf{r}(y,t)$. He uses induction over time: we suppose the reference conditions are fulfilled at time t. We will now show that there is a procedure by which we can obtain $\mathbf{R}(x,y,t+\delta t)$ for small δt.

Indeed, let us consider:

$$\hat{\mathbf{U}}(x,y,t+\delta t)=\mathbf{r}(y,t+\delta t)-\mathbf{R}(x,y,t) \tag{B.34}$$

In general, of course, the second and the third reference conditions do not hold for this displacement field $\hat{\mathbf{U}}(x,y,t+\delta t)$. This is because the reference shape at the previous time step is used in its definition.

We can form the vectors:

$$\begin{aligned}\mathbf{C}(x)&=\frac{1}{M_{\delta V(x)}}\int_{\delta V(x)}\mathbf{U}(x,z,t+\delta t)dm(z)\\ \mathbf{\Theta}(x)&=\frac{1}{2M_{\delta V(x)}}\int_{\delta V(x)}\nabla\times\mathbf{U}(x,z,t+\delta t)dm(z)\end{aligned} \tag{B.35}$$

Here, $\mathbf{\Theta}(x)$ can be considered the average rotation angle and $\mathbf{C}(x)$ can be considered the average translational displacement we can use to "correct" the position of $\mathbf{R}(x,y,t)$ to obtain $\mathbf{R}(x,y,t+\delta t)$. So that for $\mathbf{R}(x,y,t+\delta t)$, the second and the third reference conditions will hold.

Unfortunately, Smith's proof of existence and uniqueness is somewhat incomplete. Smith failed to show explicitly that we could obtain a reference shape for any time $t: t_0 \le t \le T, T-t_0>0$. Being able to advance over a (infinitesimally) small δt starting from t_0 does not guarantee that we will be able to ever reach T: formally speaking, we may

"get stuck" with fast decreasing δt, for example $\delta t_n = \frac{1}{2^n}$, so that T is never reached; in this example, it will be the case if $t_0 = 0 \wedge T > 2$, since $\sum_{n=1}^{\infty} \frac{1}{2^n} = 2$.

However, this is only a formal problem. It is easy to extend Smith's proof to show that indeed we can always reach any T. We suppose the reference conditions are fulfilled at time t_0, and we want to show that we can always reach $T : t_0 \le t \le T \wedge T - t_0 > 0$. We consider $\delta t = \frac{T - t_0}{N}$ with an arbitrary integer N. We can make δt arbitrarily small by increasing N. We can also make sure that the displacement field is uniformly small, that is

$$\left|\mathbf{U}(x, y, t_0 + n\delta t)\right| << 1, \forall n \in \{0,1,...,N\}, \forall (x, y) \in \mathrm{B}^2 \qquad \text{(B.36)}$$

We assume it is always possible to do this, since, physically, inability to obtain uniformly small displacement field at any N would mean that we are dealing with a physically abnormal deformation (discontinuous displacement field).

We can now apply Smith's procedure N times to show that we can advance from t_0 to T using this small time step δt that we have just obtained.

Decomposition of the reference shape

After all the discussion about the reference shape $\mathbf{R}(x, y, t)$ and its composition, it is not difficult to find its general form.

Indeed, let us define the following quantity:

$$\mathbf{C}(x,t) = \frac{1}{M_{\delta V(x)}} \int_{\delta V(x)} \mathbf{r}(z,t)\, dm(z)$$

$$\mathbf{C}(x,0) = \frac{1}{M_{\delta V(x)}} \int_{\delta V(x)} \mathbf{r}(z,0)\, dm(z) = \frac{1}{M_{\delta V(x)}} \int_{\delta V(x)} \mathbf{N}(z)\, dm(z) \qquad \text{(B.37)}$$

It is the centre of mass of the set $\{\mathbf{r}(z,t) : z \in \delta V(x)\}$. We must note here that Smith has a typographical error in this definition, he used $\mathbf{R}(x,z,t)$ instead of $\mathbf{r}(z,t)$ (it is easy to see why if one looks at the general form of $\mathbf{R}(x,y,t)$ below and consider a limiting case when $\delta \to 0$).

We can now write down the general form of the reference frame $\mathbf{R}(x,y,t)$:

$$\mathbf{R}(x,y,t)=\mathbf{C}(x,t)+\Omega[\Theta(x,t)](\mathbf{N}(y)-\mathbf{C}(x,0)) \qquad \text{(B.38)}$$

Where $\Omega[\Theta(x,t)]$ is the rotational transformation, rotating through the angle $\Theta(x,t)$.

In this expression, $\mathbf{C}(x,t)$ is used for translational alignment and $\Omega[\Theta(x,t)]$ is used for rotational alignment. We can also note that for any fixed x and time t, the set $\{\mathbf{R}(x,y,t): y\in \mathrm{B}\}$ has the same metric properties as the set $\{\mathbf{N}(y): y\in \mathrm{B}\}$, as it should be, of course.

Projection of the base components and displacement field

In this section Smith derives a number of formulae to facilitate the practical calculation of $\mathbf{Q}(x,y,t)$, defined as follows:

$$\mathbf{Q}(x,y,t)=\Omega[\Theta(x,t)](\mathbf{N}(y)-\mathbf{C}(x,0)) \qquad \text{(B.39)}$$

He shows that it is possible to calculate $\mathbf{Q}(x,y,t)$ everywhere in the six-dimensional domain B^2 knowing its value only at the diagonal $\mathbf{Q}(z,z,t)$, where $z\in \mathrm{B}$.

This is however only a formal improvement. The theory of Hyper-Matter itself is a special case of the six-dimensional model presented above when $\delta \to 0$. Thus, we will never need to calculate any six-dimensional quantities beyond the (infinitesimally) small six-dimensional neighbourhood of the diagonal points $(x,x)\in \mathrm{B}^2$.

In the light of that, this section has little practical value and we will not discuss it.

The six-dimensional strain tensor

In this section, Smith introduces the strain tensor. It is formally six-dimensional (since it is defined as a function of the six-dimensional displacement field), but it can be shown that in fact it is three-dimensional.

We know that in the classical case, the strain tensor can be expressed both in terms of the metric tensor (see (B.8)) and the displacement field (see (B.10) and (B.11)). In the six-dimensional case, the metric tensors are formally defined over the six-dimensional domain:

$$g_{ij}(x,y,t)=\mathbf{R}_i(x,y,t)\cdot\mathbf{R}_j(x,y,t) \qquad \text{(B.40)}$$

$$G_{ij}(y) = \mathbf{r}_i(y,t) \cdot \mathbf{r}_j(y,t)$$

But as it is easy to see, $g_{ij}(x,y,t) \equiv g_{ij}(y,t)$, that is, the metric tensor on the reference frame is independent of x. From this we can conclude that the strain tensor, which can be expressed in terms of the two metric tensors (B.40), can be expressed in the following form:

$$\varepsilon_{ij}(x,t) = \frac{1}{2}\left(\mathbf{U}_{i;j}(x,x,t) + \mathbf{U}_{j;i}(x,x,t) + \mathbf{U}_{k;i}(x,x,t) \cdot \mathbf{U}^{k}{}_{;j}(x,x,t)\right) \qquad \text{(B.41)}$$

Or, dropping the non-linear terms, as in the classical case, provided the gradients of the displacement fields are small everywhere in the δ-neighbourhood:

$$\varepsilon_{ij}(x,t) = \frac{1}{2}\left(\mathbf{U}_{i;j}(x,x,t) + \mathbf{U}_{j;i}(x,x,t)\right) \qquad \text{(B.42)}$$

The domain of linearity

In this section, Smith discusses the practical choices for the parameter δ, which is used in the six-dimensional theory to define the domain of integration in the reference conditions and the domain where the displacement gradients must be small to justify dropping non-linear terms in (B.41).

Formally, δ can assume any value $0 < \delta \leq 1$. In reality, however, only two cases are of practical interest: when $\delta = 1$ and $\delta \to 0$. Indeed, if we take a look at the reference conditions, we can see δ defines the size of the neighbourhood of a point that is used in the calculation of all local quantities at the point (strain, stress and elastic force); it defines the area in the domain where the quantities, such as the displacement field, are averaged and then used in the reference conditions. It is therefore clear that the only two distinctively different cases, when we can expect to obtain substantially different results, are when $\delta = 1$ and $\delta \to 0$. When $\delta = 1$ the whole of the domain B is used to calculate all the local quantities at the point. When $\delta \to 0$ only an infinitesimal neighbourhood of the point is used for the calculation. Any fixed value of $\delta > 0$ is similar to the case of $\delta = 1$.

Let us first consider $\delta = 1$. In this case, $\delta V(x,y) \equiv \{(x,z) : z \in \mathrm{B}\}$ and $\delta V(x) \equiv \mathrm{B}$, that is the integration domain in the reference conditions as well as the domain where the displacement gradients must be small coincides with the whole of B for every point. The reference

conditions are therefore independent of x and only depend on time; the second and third alignment reference conditions are therefore global alignment conditions.

This case corresponds to the global alignment approach, considered in the present thesis (see section 4.3.1) as a possible extension of the classical theory of elasticity to add rigid body translations and rotations. It also relates to the hybrid formulation by Terzopoulos et al [7, 12]. This global alignment approach provides a clear improvement over the classical theory of elasticity; it allows the reference frame to undergo rigid body translations and rotations. But, as we discussed in the present thesis in section 4.3.1, it still places severe limitations of the range of allowed elastic deformation (due to the requirement of globally small displacement field gradients).

We will now consider the limiting case when $\delta \to 0$. This case is of considerable practical interest. In the reference conditions, as the domain of integration shrinks to a point, the reference conditions become purely local for every point of the solid (that is, the second and the third reference conditions become the conditions for the quantities in question at a point itself and are no longer averaged over a finite neighbourhood):

$$\begin{aligned} U^i(x,x,t) &= 0 \\ \in^{ijk} U_{j;k}(x,x,t) &= 0 \end{aligned} \tag{B.43}$$

Similarly, in (B.42) the requirement of small displacement gradients is also purely local; the gradients only need to be small in an infinitesimally small neighbourhood of a point to justify dropping the non-linear term in (B.41). This significantly simplified requirement for the displacement gradients lies at the heart of the flexibility of the theory of Hyper-Matter and our own theory of FlexyMatter, and accounts for their ability to model globally large deformations. We consider this to be a highly significant advance on what one is able to achieve in classical elasticity.

Other simplifications in the six-dimensional theory when $\delta \to 0$ are also possible. Because of the second reference condition (the first equality in (B.43)), we can see that the first covariant derivative can be replaced with the first partial derivative:

$$U^i(x,x,t) = 0 \Rightarrow \tag{B.44}$$

$$U_{j;k}(x,x,t)=U_{j,k}(x,x,t)-\Gamma^{i}_{jk}(x,x)U_{i}(x,x,t)=U_{j,k}(x,x,t)$$

Therefore, the second and the third reference conditions can be simplified even further:

$$\begin{aligned} U^{i}(x,x,t)&=0 \\ \in^{ijk} U_{j,k}(x,x,t)&=0 \end{aligned} \tag{B.45}$$

Also, the expression for the strain can be simplifies considerably:

$$\varepsilon_{ij}(x,t)=\frac{1}{2}\left(U_{i,j}(x,x,t)+U_{j,i}(x,x,t)\right)=U_{i,j}(x,x,t)=U_{j,i}(x,x,t) \tag{B.46}$$

Smith refers to the parameter δ as the representative of the domain of linearity of the model, hence the name of this section. Indeed, we must place a restriction of the small gradients of the displacement field everywhere in $\delta V(x)$ to justify using the linear expression for the strain (B.42) instead of the non-linear one in (B.41).

At the end of this section, Smith states that the only case of considerable practical interest is when $\delta \to 0$. In this case, the theory can model almost arbitrary global displacements, allowing for globally large elastic deformations. He argues that anything that can be modelled by setting $\delta = 1$ can also be modelled by setting $\delta \to 0$.

Smith calls the six-dimensional theory with special case of $\delta \to 0$ the theory of Hyper-Matter. It is the subject of the remaining part of his thesis, which is considered below.

B.1.3 The Theory of Hyper-Matter

In the previous sections we have provided an overview of a six-dimensional theory, developed by Paddon and Smith and presented in Smith's PhD thesis. This theory has a parameter δ, which relates to the domain of the linearity of the theory where the principles of the linear classical theory of elasticity can be applied (most importantly, where the displacement gradients are required to be small to justify using a linear expression for the strain). We have explained in the previous section that even though δ can take any value $0 < \delta \leq 1$, only two special cases present practical interest: when $\delta = 1$ and when $\delta \to 0$.

The remaining part of Smith's thesis deals with the special case of $\delta \to 0$. A new notation is introduced specially for this case. The six-dimensional theory itself is called the theory of

Hyper-Matter. The general six-dimensional shape, represented by $\mathbf{R}(x,y,t)$, is called a hyper-body and the reference frame $\mathbf{R}(x,y,t)$ is renamed into $\mathbf{H}(x,y,t)$.

It is useful to present the reference conditions for the hyper-body (which can be obtained directly from their more generic form (B.25), (B.26) and (B.27) assuming $\delta \to 0$):

$$\begin{gathered}\mathbf{H}_i(x,y,t)\cdot\mathbf{H}_j(x,y,t)=\mathbf{N}_i(y)\cdot\mathbf{N}_j(y)\\ \mathbf{H}(y,y,t)=\mathbf{r}(y,t)\\ \in^{ijk} H_{j;k}(y,y,t)=\in^{ijk} r_{j;k}(y,t)\Leftrightarrow\\ \in^{ijk} H_{j;k}(y,y,t)=\in^{ijk} H_{j;k]}(y,y,t)\end{gathered} \tag{B.47}$$

Here, in the last reference condition, the notation $H_{j;k]}(y,y,t)$ indicates that the two spatial variables are held together during the differentiation, for example for $k = 2$:

$$\mathbf{H}_{,2]}(y,y,t)=\lim_{dy^2\to 0}\frac{\mathbf{H}(y^1,y^2+dy^2,y^3,y^1,y^2+dy^2,y^3t)}{dy^2} \tag{B.48}$$

The displacement field $\mathbf{U}(x,y,t)$ is, as before:

$$\begin{gathered}\mathbf{U}(x,y,t)=\mathbf{r}(y,t)-\mathbf{H}(x,y,t)=\\ =[ref\ condition\ 2]=\mathbf{H}(y,y,t)-\mathbf{H}(x,y,t)\end{gathered} \tag{B.49}$$

We can also formulate the second and the third reference conditions in terms of the displacement field:

$$\begin{gathered}\mathbf{U}(y,y,t)=\mathbf{0}\\ \in^{ijk} U_{j;k}(y,y,t)=0\end{gathered} \tag{B.50}$$

These are the alignment conditions for the deforming solid and its reference frame at each point and at all times (the first condition describes the translational alignment and the second - the rotational one).

We will now briefly depart from following Smith's thesis and compare the results we have so far with the classical case. As we can see in (B.16), the equations of motion describing the evolution of the deforming solid contain the terms with the stress at each point of the

solid. This stress is linearly expressed through the strain (via the constitutive equations), which, in its turn, is expressed via the gradients of the displacement field at the point:

$$\varepsilon_{ij}(x,t)=\frac{1}{2}\left(U_{i;j}(x,t)+U_{j;i}(x,t)\right)$$
$$\mathbf{U}(x,t)=\mathbf{r}(x,t)-\mathbf{N}(x) \qquad \text{(B.51)}$$

Where $\mathbf{N}(x)$ is the reference frame, being the solid at the start of the experiment.

As we have shown, this classical scheme is not adequate for modelling globally large deformations primarily due to the local (rotational) misalignments of the deforming solid and the reference frame (more precisely, the expression for the stress in (B.51) contains an error term on the right due to the displacement field containing a rotational component).

We can see that the theory of Hyper-Matter addresses this misalignment problem in the following way. It introduces a hyper-body $\mathbf{H}(x,y,t)$ (the generalised reference frame) and three reference conditions this hyper-body must satisfy (B.47) for all points and at all times. Then a generalised displacement field (B.49) is introduced, from which an alternative expression for the strain can be derived (B.46). Or, to summarise:

$$\varepsilon_{ij}(x,t)=\frac{1}{2}\left(U_{i,j}(x,x,t)+U_{j,i}(x,x,t)\right)=U_{i,j}(x,x,t)=U_{j,i}(x,x,t)$$
$$\mathbf{U}(x,y,t)=\mathbf{r}(y,t)-\mathbf{H}(x,y,t) \qquad \text{(B.52)}$$

Plus of course the three reference conditions (B.47).

After we obtained the expression (B.52) for the strain free from the misalignment errors, we can continue the procedure as in the classical scheme, and obtain the new equations of motion (B.16).

We can now compare the theory of Hyper-Matter directly with the classical theory of elasticity ((B.52) and (B.51)). We introduce a generalised reference frame $\mathbf{H}(x,y,t)$ instead of the fixed one $\mathbf{N}(x)$ as in the classical case, together with the reference conditions (B.47), which uniquely define $\mathbf{H}(x,y,t)$. Then we introduce a generalised displacement field (B.52), defined over a six-dimensional domain, as opposed to the classical three-dimensional one (B.51). From this, as in the classical case, we obtain an expression for the strain which is free from alignment errors but which is defined over the three-dimensional

domain. After this, the procedure to obtain the equations of motions is the same as in the classical theory of elasticity.

We can also note that, in addition to obtaining the expression for the strain that is free of alignment errors, the theory of Hyper-Matter also significantly simplified the expression for the strain (by replacing the covariant derivative with the partial one and simplifying the strain expression itself). Also, the covariant derivative can be replaced with the partial derivative and expressions for the strain. Therefore, stress and the elastic force are simplified (due to the third reference condition).

We have now provided a complete practical description of the theory of Hyper-Matter, even though we have only just started the chapter "The Theory of Hyper-Matter" from Smith's thesis. The rest of this chapter provides a collection of very theoretical discussions on various aspects of the generalised reference frame $\mathbf{H}(x,y,t)$ with little practical value (except perhaps for the last section, "Energy Balance", where Smith discusses energy balance of the hyper-body $\mathbf{H}(x,y,t)$). In addition, as Smith considers all the relevant entities, such as the stress, elastic force and so on, to be formally defined over the six-dimensional material domain and, as usual, makes little effort to explain what he is actually trying to achieve, the remaining part of this chapter looks complex and confusing. As these sections are not material in understanding the theory of Hyper-Matter, we only give a brief overview of them, without too much technical detail:

- In the "Transformation System" section, Smith derives a number of classical formulae for the tensor transformations between the material coordinate system and the world coordinate system.

- In the "Physical Representation of the Elastic Force" section Smith considers the six-dimensional extension of the classical three-dimensional elastic force $\mathcal{E}^i(x,y,t)$, and looks at the physical meaning of this extension outside the diagonal sub-domain (when x and y are held together). Smith obtains an entirely expected result: only at the diagonal sub-domain does the formally extended elastic force make physical sense, whilst outside it the elastic force is of the correct magnitude but misaligned rotationally and as such makes little physical sense. Smith concludes that it only makes sense to consider $\mathcal{E}^i(x,y,t)$ at the diagonal sub-domain.

- In the next two sections named "XY-Reducibility" and "Colateral Derivative" Smith introduces the notions of XY-reducibility of an entity defined over a six-dimensional domain and its colateral derivative. These two notions are used for purely theoretical manipulations with very little practical value and we will not cover them in any detail.

- In the section "Dynamics" Smith shows that the total sum of the moments of the elastic forces inside the diagonal sub domain (all the entities in question, such as, the stress, elastic force and so on are formally extended into six material dimensions) of the hyper-body is zero (he does it by using the Divergence theorem), provided the local displacement gradients are small. When the displacement gradients can no longer be considered small locally, Smith argues, this total sum of the moments is no longer zero. From this, Smith makes the following conclusion: the point elastic forces must be treated with care when dealing with rotational motion of the deforming body.

- In the last section "Energy Balance" Smith proves that the total energy of the system is preserved throughout the motion. That is, $K + P = const$, where K is the total kinetic energy of the system and P is the total strain energy (or the total elastic potential energy) of the system. As usual, he considers all the quantities as being formally defined over the six-dimensional material domain; however, as before, in all the expressions they are restricted to the diagonal sub domain.

B.1.4 Material elements

In the chapter called "Material Elements" Smith has developed a model for the treatment of material elements – the elementary pieces of material treated as whole units, obtained after the process of discretization. This model includes a system of symbolic notation to describe the material elements with their diverse range of possible boundary configurations and a system by which the elastic force acting on any material element can be practically estimated.

Understanding the model requires a good knowledge of set theory; the presentation style is as usual terse, and this model is not material to the theory of Hyper-Matter.

B.2 Conclusion

The theory of Hyper-Matter, developed by Smith and Paddon [89, 90], was the first comprehensive and rigorous modelling theory based on the classical theory of elasticity. It was distinctly different from any other modelling theory developed by the research community over the last 20 years, which are mostly based on the energy methods.

Despite its novel approach, the theory of Hyper-Matter had a poor reception. There are a number of reasons for this, including:

- Poor presentation style. The theory itself is presented in a very terse style, with no examples and little background explanations for the key ideas. The merits of presenting a generic, full-blown six-dimensional theory with a parameter δ (relating to the domain of linearity) first, followed by the theory of Hyper-Matter as a special case of $\delta \to 0$, are questionable. The six-dimensional theory, which is presented first, is more complex and with no examples and little explanation is difficult to grasp. The theory of Hyper-Matter is then presented as a special case relying on the reader to have a good understanding of the more general six-dimensional theory.

- The actual need for a full six-dimensional theory is questionable. As we have shown in this thesis, the same key ideas can be turned into a much simpler theory (the theory of FlexyMatter). Indeed, as we have noted in this appendix, not only does the need to deal with six-dimensional quantities require the (formal) expansion of the usually three-dimensional quantities (stress, strain, the displacement field etc) into six-dimensions and the introduction of new operations (such as taking a derivative collaterally), but also in the majority of cases the expanded six-dimensional quantities are only considered in the diagonal (three-dimensional) sub domain of B^2. We have also shown that, even when using the six-dimensional expansion, it is possible to use the expansion very sparingly (namely, only use the six-dimensional displacement field), without formally expanding other quantities into six-dimensions.

The theory of Hyper-Matter was a very important step on the way of creating a better modelling theory; without it, our own thesis may never have been conceived. In the theory of Hyper-Matter Smith and Paddon showed that it is possible to create a rigorous and flexible modelling theory based firmly on the classical theory of elasticity.

The theory of Hyper-Matter however suffered from a number of shortcomings, which we have outlined in the present thesis and which negatively affected a wide acceptance of the theory (and the key ideas) by the research community. One of the aims of our own theory of FlexyMatter was to do a complete overhaul of the theory of Hyper-Matter to try and make it much more acceptable and appealing to a wide audience of researchers, engineers and computer animators. We hope that this will help the key ideas, which were first presented in Smith's thesis and are now redeveloped in our theory of FlexyMatter, be used in the future advanced modelling techniques and get the acceptance and the recognition in the research community that they rightly deserve.

Bibliography

Articles:

1980 – 1985:

1. **Kawabata, S.,** "The Standardization and Analysis of Hand Evaluation", The Textile Machinery Society of Japan, Osaka, 1980

2. **Ohashi, T., Tokoro., M., Uchiki, T.,** "Collision detection in motion simulation", Comput. & Graphics, V. 7, No. 3-4, 1983

3. **Armstrong, W., W., Green, M.,** "The dynamics of articulated rigid bodies for purposes of animation", Proc. Graphics Interface'85, Montreal, Canada, 1985

4. **Wilhelms, J., Barsky, B., A.,** "Using dynamic analysis to animate articulated bodies such as humans and robots", Proc. Graphics Interface'85, Montreal, Canada, 1985

1986:

5. **Weil, J.,** "The synthesis of cloth objects", Proc. Siggraph'86, Computer Graphics, V. 20, No. 4, 1986

6. **Feynman, C., R.,** "Modelling the appearance of cloth", MSc Thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1986

1987:

7. **Terzopoulos, D., Platt, J., Barr, A., Fleischer, K.,** "Elastically deformable models", Computer Graphics, V. 21, No. 4, 205-214, 1987

8. **Cohen, M., F., Isaacs, P., M.,** "Controlling dynamic simulation with kinematic constraints, behaviour functions and inverse dynamics", Computer Graphics, V. 21, No. 4, 1987

1988:

9. **Miller, G., S., P.,** "The motion dynamics of snakes and worms", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

10. **Kass, M., Witkin, A.,** "Spacetime constraints", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

11. **Barr, A., H., Platt, J., C.,** "Constraint method for flexible models", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

12. **Fleischer, K., Terzopoulos, D.,** "Modelling inelastic deformation: viscoelasticity, plasticity, fracture", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

13. **Barr, A., H., Barzel, R.,** "A modelling system based on dynamic constraints", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

14. **Moore, M., Wilhelms, J.,** "Collision detection and response for computer animation", Proc. Siggraph'88, Computer Graphics, V. 22, No. 4, 1988

15. **Magnenat-Thalmann, N., Laperriere, R., Thalmann, D.,** "Joint-dependant local deformations for hand animation and object grasping", Proc. Graphics Interface'88, Edmonton, 1988

1989:

16. **Chadwick, J., E., Haumann, D., R., Parent., R., E.,** "Layered construction for deformable characters", Proc. Siggraph'89, Computer Graphics, V. 23, No. 3, 1989

17. **Terzopoulos, D., Szeleski, R.,** "From splines to fractals", Proc. Siggraph'89, Computer Graphics, V. 23, No. 3, 1989

1990:

18. **Aono, M.,** "Wrinkle propagation model for cloth", Proc. Computer Graphics International'90, Springer, Tokyo, 1990

19. **Witkin, A., Gleicher, M., Welch, W.,** "Interactive dynamics", Computer Graphics, 1990

1991:

20. **Lafleur, B., Magnenat-Thalmann, N., Thalmann, D.,** "Cloth animation with self-collision", Proc. IEIP Conference on modelling in Computer Graphics'91, Springer, New York, 1991

21. **Gascuel, M.-P., Verroust, A., Puech, C.,** "A modelling system for complex deformable bodies suited to animation and collision detection", The Journal of Visualisation and Computer Animation, V. 2, 82-91, 1991

1992:

22. **Carignan, M., Yang, Y., Magnenat-Thalmann, N., Thalmann, D.,** "Dressing animated synthetic actors with complex deformable clothes", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 99–104, 1992

23. **Terzopoulos, D., Metaxas, D.,** "Dynamic deformation of solid primitives with constraints", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 309-312, 1992

24. **Cohen, M.,** "Interactive spacetime control for animation", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 293-302, 1992

25. **Welch, W., Witkin, A.,** "Variational surface modelling", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 157-166, 1992

26. **Hsu, W., M.,** "Direct manipulation of free-form deformation", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 177-184, 1992

27. **Baraff, D., Witkin, A.,** "Dynamic simulation of non-penetrating flexible bodies", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 303-308, 1992

28. **Van Overveld, C., W., A., M., Van Loon, E.,** "Hanging cloth and dangling rods: a unified approach to constraints in computer animation", The Journal of Visualisation and Computer Animation, V. 3, 45-59, 1992

29. **Okabe, H., Imaoka, H., Tohima, T., Hiwaya, H.,** "Three dimensional apparel CAD system", Proc. Siggraph'92, Computer Graphics, V. 26, No. 2, 105-110, 1992

1993:

30. **Gascuel, M.-P.,** "An implicit formulation for precise contact modelling between flexible solids", Proc. Siggraph'93, Annual Conference Series, 1993

31. **Snyder, J., M., Woodbury, A., R., Fleischer, K., Currin, B., Barr, A.,** "Interval methods for multi-point collisions between time-dependent curved surfaces", Proc. Siggraph'93, Annual Conference Series, 1993

32. **Ngo, J., T., Marks, J.,** "Spacetime constraints revisited", Proc. Siggraph'93, Annual Conference Series, 343-350, 1993

33. **Lin, M., C., Manocha, D.,** "Interference detection between curved objects for computer animation", Proc. Computer Animation'93, 43-57, 1993

34. **Ling, L., Damodaran, M., Gay, R.,** "A quasi-steady force model for animating cloth motion", Graphics, Design and Visualisation, North Holland, 357-363, 1993

35. **Dhande, S., G., Rao, P., V., M., Tavakkoli, S., Moore, C., L.,** "Geometric modelling of draped fabric surfaces", Graphics, Design and Visualisation, North Holland, 349-356, 1993

1994:

36. **Gudukbay, U., Ozguc, B.,** "Animation of deformable models", Computer Aided Design, V. 26, No. 12, 868-875, 1994

37. **Baraff, D.,** "Fast contact force computation for non-penetrating rigid bodies", Proc. Siggraph'94, Annual Conference Series, 1994

38. **Breen, D., E., House, D., H., Wozny, M., J.,** "Predicting the drape of woven cloth using interacting particles", Proc. Siggraph'94, Annual Conference Series, 1994

39. **Breen, D., E., House, D., H., Wozny, M., J.,** "A particle-based model for simulating the draping of woven cloth", Textile Research Journal, V. 64, No. 11, 663-685, 1994

40. **Witkin, A., P., Heckbert, P., S.,** "Using particles to sample and control implicit surfaces", Proc. Siggraph'94, Annual Conference Series, 1994

41. **Tu, X., Terzopoulos, D.,** "Artificial fishes: physics, locomotion, perception, behaviour", Proc. Siggraph'94, Annual Conference Series, 1994

42. **Liu, Z., Gortler, S., J., Cohen, M., F.,** "Hierarchical spacetime control", Proc. Siggraph'94, Annual Conference Series, 1994

43. **Volino, P., Magnenat-Thalmann, N.,** "Efficient self-collision detection on smoothly discretised surface animations using geometrical shape regularity", Eurographics'94, V. 13, No. 3, 155-166, 1994

1995:

44. **Grzeszczuk, R., Terzopoulos, D.,** "Automated learning of muscle-actuated locomotion through control abstraction", Proc. Siggraph'95, Annual Conference Series, 1995

45. **Lee, Y., Terzopoulos, D., Waters, K.,** "Realistic modelling for facial animation", Proc. Siggraph'95, Annual Conference Series, 1995

46. **Volino, P., Courchesne, M., Magnenat-Thalmann, N.,** "Versatile and efficient techniques for simulating cloth and other deformable objects", Proc. Siggraph'95, Annual Conference Series, 1995

47. **Gobetti, E., Balaguer, J.-F.,** "An integrated environment to visually contract 3D animations", Proc. Siggraph'95, Annual Conference Series, 1995

48. **Harada, H., Witkin, A., Baraff, D.,** "Interactive physically based manipulation of descrete/continuus models", Proc. Siggraph'95, Annual Conference Series, 1995

49. **Desbrun, M., Gascuel, M.-P.,** "Animating soft substances with implicit surfaces", Proc. Siggraph'95, Annual Conference Series, 1995

50. **Hodgins, J., K., Wooten W., L., Brogan, D., C., O'Brien, J., F.,** "Animating human athletics", Proc. Siggraph'95, Annual Conference Series, 1995

51. **Fekete, J.-D., Bizouarn, E., Cournarie, E., Galas, T., Taillefer, F.,** "TicTacToon: a paperless system for professional 2D animation", Proc. Siggraph'95, Annual Conference Series, 1995

52. **Ng, H., N., Grimsdale, R., L., Allen, W., G.,** "A system for modelling and visualization of cloth material", Comput.&Graphics, V. 19, No. 3, 423-430, 1995

53. **Provot, X.,** "Deformation constraints in a mass-spring model to describe rigid cloth behaviour", Proc. Graphics Interface'95, 147-157, 1995

54. **Keller, H., Stolz, H., Andreas, Z., Braunl, T.,** "Virtual mechanics simulation and animation of rigid body systems with AERO", Technical article, Simulation, V. 65, No. 1, 74-78, 1995

55. **Sakaguchi, Y., Minoh, M., Ikeda, K.,** "PARTY: a numerical calculation method for a dynamically deformable cloth model", Systems and Computers in Japan, V. 26, No. 8, 1995

56. **Baraff, D.,** "Interactive simulation of solid rigid bodies", IEEE Computer Graphics and Applications, V. 15, No. 3, 63-75, 1995

57. **Shen, J., Thalmann, D.,** "Interactive shape design using meatballs and splines", Proc. Eurographics on Implicit Surfaces'95, 187-196, 1995

1996:

58. **Eberhardt, B., Weber, A., Strasser, W.,** "A fast, flexible, particle-system model for cloth draping", IEEE Computer Graphics and Applications, September, 1996, 52-59

59. **Milenkovic, V., J.,** "Position-based physics: simulating the motion of many highly interacting spheres and polyhedra", Proc. Siggraph'96, Annual Conference Series, 1996

60. **Rose, C., Guenter, B., Bodenheimer, B., Cohen, M., F.,** "Efficient generation of motion transitions using spacetime constraints", Proc. Siggraph'96, Annual Conference Series, 1996

61. **Mark, W., R., Randolph, S., C., Finch, M., Van Verth, J., M., Taylor II, R., M.,** "Adding force feedback to graphics systems: issues and solutions", Proc. Siggraph'96, Annual Conference Series, 1996

62. **Baraff, D.,** "Linear-time dynamics using Lagrange multipliers", Proc. Siggraph'96, Annual Conference Series, 1996

63. **Magnenat-Thalmann, N., Volino, P., Yang, Y.,** "3D clothes and fashion show", 4th International Conference in Central Europe on Computer Graphics and Visualization'96, 2-20, 1996

64. **Magnenat-Thalmann, N., Volino, P., Jianhua, S., Thalmann, D.,** "An evolving system for simulating clothes on virtual actors", IEEE Computer Graphics and Applications, September, 1996, 42-51

65. **Qin, H., Terzopoulos, D.,** "D-NURBS: a physically based framework for geometric design", IEEE Computer Graphics and Applications, September, 1996, 42-51

66. **Mudur, S., P.,** "Physically based modelling for multimedia virtual environments", IETE Technical Review, V. 13, No. 6, 303-315, 1996

67. **Ling, L., Damoradan, M., Gay., R., K., L.,** "A model for animating the motion of cloth", Comput.&Graphics, V. 20, No. 1, 137-156, 1996

1997:

68. **Scheepers, F., Parent, R., E., Carlson, W., E., May, S.,** "Anatomy-based modelling of the human musculature", Proc. Siggraph'97, Annual Conference Series, 1997

69. **Wilhelms, J., Van Gelder, A.,** "Anatomically based modelling", Proc. Siggraph'97, Annual Conference Series, 1997

70. **Breen, D., E., Eischen, J., W., Kass, M., Magnenat-Thalmann, N., Vecchione, M.,** "Can we get there from here?: Current challenges in cloth design, modelling and animation", Panel, Proc. Siggraph'97, Annual Conference Series, 1997

1998:

71. **Singh, K., Fiume, E.,** "Wires: A geometric deformation technique", Proc. Siggraph'98, Annual Conference Series, 1998

72. **Baraff, D., Witkin, A.,** "Large steps in cloth simulation", Proc. Siggraph'98, Annual Conference Series, 1998

73. **Volino, P., Magnenat-Thalmann, N.,** "Interactive cloth simulation: Problems and solution", on-line article, http://miralabwww.unige.ch/ACTICLES/JWS97-B.html, 1998

1999:

74. **Remion, Y., Nourrit, J.-M., Gillard, D.,** "Dynamic animation of spline-like objects", 1999

75. **Desbrun, M., Schroder, P., Barr, A.,** "Interactive animation of structured deformable objects", Proc. Graphics Interface'99, 1999

76. **Kang, Y.-M., Choi, J.-H., Cho, H.-G.,** "Real-time animation technique for flexible and thin objects", 1999

77. **You, L., Zhang, J., J., Comninos, P.,** "Cloth deformation modelling using a plate bending model", 1999

78. **Remion, Y., Nourrit, J.-M., Nocent, O.,** "Dynamic animation of n-dimensional deformable objects", 1999

79. **James, D., L., Pai, D., K.,** "ArtDefo: Accurate real time deformable objects", Proc. Siggraph'99, Annual Conference Series, 1999

80. **Funge, J., Tu, X., Terzopoulos, D.,** "Cognitive Modelling: Knowledge, reasoning and planning for intelligent characters", Proc. Siggraph'99, Annual Conference Series, 1999

81. **Lee, J., Shin, S., Y.,** "A hierarchical approach to interactive motion editing for human-like figures", Proc. Siggraph'99, Annual Conference Series, 1999

82. **Usoh, M., Arthur, K., Whitton, M., C., Bastos, R., Steed, A., Slater, M., Brooks, F., P.,** "Walking > Walking-in-space > Flying, in virtual environments", Proc. Siggraph'99, Annual Conference Series, 1999

83. **O'Brien, J., F., Hodgins, J., K.,** "Graphical modelling and animation of brittle fracture", Proc. Siggraph'99, Annual Conference Series, 1999

2000:

84. **Smith, J., Witkin, A., Baraff, D.,** "Fast and controllable simulation of the shattering of brittle objects", Proc. Siggraph'00, Annual Conference Series, 2000

85. **Popovic, J., Seitz, S., M., Erdmann, M., Popovoc, Z., Witkin, A.,** "Interactive manipulation of rigid body simulations", Proc. Siggraph'00, Annual Conference Series, 2000

86. **Mirtich, B.,** "Timewarp rigid body simulation", Proc. Siggraph'00, Annual Conference Series, 2000

87. **Alexa, M., Cohen-Or, D., Levin, D.,** "As-rigid-as-possible shape interpolation", Proc. Siggraph'00, Annual Conference Series, 2000

88. **Lewis, J., P., Cordner, M., Fong, N.,** "Pose space deformation: A unified approach to shape interpolation and skeleton-driven deformation", Proc. Siggraph'00, Annual Conference Series, 2000

Books and other references:

89. **Smith, J.,** "Computer Modelling of the General Dynamics of Deformable Bodies", PhD Thesis, Department of Computer Science, University of Bristol, 1992

90. **Smith, J., Paddon, D., J.,** "Hyper-Matter", Kaleidoscope, Technical Report, KJ-92-01 (CSTR-92-38), Department of Computer Science, University of Bristol, 1992

91. **Flugge, W.,** "Continuum Mechanics and Tensor Analysis", Springer-Verlag, 1972

92. **Goldstein, H.,** "Classical Mechanics", Addison-Wesley, Reading, MA, 1950

93. **Chung, T., J.,** "Applied Continuum Mechanics", Cambridge University Press, 1996

94. **Timoshenko, S., P., Goodier, J., N.,** "Theory of Elasticity", McGraw-Hill International Editions, Third Edition, 1970

95. **Jaunzemis, W.,** "Continuum Mechanics", Macmillan, 1967

96. **Truesdell, C.,** "The Elements of Continuum Mechanics", Springer-Verlag, 1966

97. **Bird, R., B., Armstrong, R., C., Hassager, O.,** "Dynamics of Polymeric Fluids", Volume 1: Fluid Mechanics, John Wiley & Sons, 1977

98. **Vekstein, G., E.,** "Physics of Continuous Media: A Collection of Problems with Solutions for Physics Students", Adam Hilger, 1992

99. **Roberts, A., J.,** "A One-Dimensional Introduction to Continuum Mechanics", World Scientific, 1994

100. **Press, W., H., Teukolsky, S., A., Vetterling, W., T., Flannery, B., P.,** "Numerical Recipes in C", The 101. Art of Scientific Computing, Second Edition, Cambridge University Press, 1999

102. **Eglit, M., E., Hodges, D., H.,** "Continuum mechanics via problems and exercises", Part I & II, World Scientific Publishing Co. Pte. Ltd, 1996

103. **Lovelock, D., Rund, H.,** "Tensors, Differential Forms, and Variational Principles", Dover Publications, Inc., New York, 1989

104. **Borisenko, A., I., Tarapov, I., E.,** "Vector and Tensor Analysis with Applications", Dover Publications, Inc., New York, 1979

105. **Woo, M., Neider, J., Davis, T., Shreiner, D.,** "OpenGL Programming Guide", Third Edition, Addison-Wesley, 1999